

003525, 00404

POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE OF  
MODULATING VARIOUS RESPONSES

## TABLE OF CONTENTS

### Title

### 5 Field Of The Invention

### Background

### Summary Of The Invention

10

### Brief Description Of The Inventions

- I. Sequence Inventions
- II. Differential Expression Inventions
- III. Phenotypic Inventions

15

### How The Inventions Reveal How Genes, Gene Components and Products Function

- I. Experimental Results Reveal Many Facets Of A Single Gene
  - I.A. Intro: Functions And Characteristics Of Gene Components and Products Are Identified
    - 20 I.A.1. Functions Of Coding Sequences Revealed By The Ceres Genomic Engine
      - A.1.a. Sequence Similarity To Proteins Of Known Function Can Be Used To Associate Biochemical Activities And Molecular Interaction To The Proteins Of The Invention
        - 25 1. Presence Of Amino Acid Motifs Indicates Biological Function
        - 2. Related Amino Acid Sequences Share Similar Biological Functions
      - A.1.b. Differential Expression Results Explain In Which Cellular Responses The Proteins Of The Invention Are Involved
        - 30 1. The Transcript Levels Of A Protein Over Time In Response To A Stimuli Are Revealed By Transcriptional Analyses Over Many Experiments.

2. The Transcript Levels Of A Protein Over Different Developmental Stages Can Be Identified By Transcriptional Analyses Over Many Experiments
3. Proteins That Are Common In A Number Of Similar Responses Can Be Identified By Transcriptional Analyses Over A Number Of Experiments
4. Proteins That Are Common To Disparate Responses Can Be Identified By Transcriptional Analyses Over A Number Of Experiments

A.1.c. Observations Of Phenotypic Changes Show What Physiological Consequences Applicants' Proteins Can Produce

I.A.2. Specificity Of Regulatory Sequence Are Uncovered By The Ceres Genomic Engine

A.2.a. Differential Expression Results Explain Which External Or Internal Stimuli Trigger The Regulatory Sequences

## II. Experimental Results Also Reveal Pathways Or Networks Of Genes

II.A. Genes Whose Transcription Are Well Coordinated Generally Act Together To Produce Proteins That Participate In The Same Pathway Or Network

II.A.1. Calculating The Correlation Coefficient Between Pairs Of Genes Based On The Differential Expression Data

II.A.2. The Complete Linkage Analyses Of Differential Identity Genes With Similar Pattern Of Transcription

II.A.3. The Nearest Neighbor Analyses Of Differential Group Genes With Correlated But Dissimilar Transcription Profiles

## III. Experimental Results Reveal The Function And Characteristic Genes, Pathways And Networks

III.A. Linking Biochemical Or Metabolic Activities Of One Protein In A Cluster To The Other Proteins In The Same Microarray Cluster

- III.B. Using Differential Expression Data To Determine When The Genes And Pathways Are Active
- III.C. Using Phenotype Data To Determine When Genes And Pathways Are Active

- 5 IV. Experimental Results Provide An Understanding Of Genes, Pathways and Networks In Many Plant Species
  - IV.A. Intro: Using Protein Sequence Similarity To Assign Function From Arabidopsis Proteins To Proteins Of Other Species
    - 10 IV.A.1. Conserved Sequences Between Proteins Of Different Species Give Rise To A Signature Sequence
    - IV.A.2. Linking Signature Sequences To Conservation Of Structural Characteristics
    - IV.A.3. Linking Signature Sequences To Conservation Of Biochemical Activities And Molecular Interactions
    - 15 IV.A.4. Linking Signature Sequences To Conservation Of Cellular Responses Or Activities
    - IV.A.5. Linking Signature Sequences To Conservation Of Phenotypic Effects

#### **Applications of the Inventions**

- 20 I. Organ Affecting Genes, Gene Components, Products (Including Differentiation Function)
  - I.A. Root Genes, Gene Components And Products
    - I.A.1. Root Genes, Gene Components And Products
    - I.A.2. Root Hair Genes, Gene Components And Products
  - 25 I.B. Leaf Genes, Gene Components And Products
    - I.B.1. Leaf Genes, Gene Components And Products
    - I.B.2. Trichome Genes And Gene Components
    - I.B.3. Chloroplast Genes And Gene Components
  - 30 I.C. Reproduction Genes, Gene Components And Products
    - I.C.1. Reproduction Genes, Gene Components And Products



I.C.2. Ovule Genes, Gene Components And Products

I.C.3. Seed And Fruit Development Genes, Gene Components And Products

5

I.D. Development Genes, Gene Components And Products

I.D.1. Imbibition and Germination Responsive Genes, Gene Components And Products

I.D.2. Early Seedling Phase Genes, Gene Components And Products

I.D.3. Size and Stature Genes, Gene Components And Products

I.D.4. Meristem Genes, Gene Components And Products

I.D.5. Vegetative-Phase Specific Responsive Genes, Gene Components And Products

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

II. Hormones Responsive Genes, Gene Components And Products

II.A. Abscissic Acid Responsive Genes, Gene Components And Products

II.B. Auxin Responsive Genes, Gene Components And Products

II.C. Brassinosteroid Responsive Genes, Gene Components And Products

II.D. Cytokinin Responsive Genes, Gene Components And Products

II.E. Gibberellic Acid Responsive Genes, Gene Components And Products

III. Metabolism Affecting Genes, Gene Components And Products

III.A. Nitrogen Responsive Genes, Gene Components And Products

III.B. Circadian Rhythm Responsive Genes, Gene Components And Products

III.C. Blue Light (Phototropism) Responsive Genes, Gene Components And Products

III.D. CO<sub>2</sub> Responsive Genes, Gene Components And Products

III.E. Mitochondria Electron Transport (Respiration) Genes, Gene Components And Products

III.F. Protein Degradation Genes, Gene Components And Products

III.G. Carotenogenesis Responsive Genes, Gene Components And Products

- IV. Viability Genes, Gene Components And Products
- IV.A. Viability Genes, Gene Components And Products
- IV.B. Histone Deacetylase (Axel) Responsive Genes, Gene Components And Products
- 5 V. Stress Responsive Genes, Gene Components And Products
- V.A. Cold Responsive Genes, Gene Components And Products
- V.B. Heat Responsive Genes, Gene Components And Products
- V.C. Drought Responsive Genes, Gene Components And Products
- V.D. Wounding Responsive Genes, Gene Components And Products
- 10 V.E. Methyl Jasmonate Responsive Genes, Gene Components And Products
- V.F. Reactive Oxygen Responsive Genes, Gene Components And H<sub>2</sub>O<sub>2</sub> Products
- V.G. Salicylic Acid Responsive Genes, Gene Components And Products
- V.H. Nitric Oxide Responsive Genes, Gene Components And Products
- V.I. Osmotic Stress Responsive Genes, Gene Components And Products
- V.J. Aluminum Responsive Genes, Gene Components And Products
- V.K. Cadmium Responsive Genes, Gene Components And Products
- V.L. Disease Responsive Genes, Gene Components And Products
- V.M. Defense (lcl2) Responsive Genes, Gene Components And Products
- V.N. Iron Responsive Genes, Gene Components And Products
- 20 V.O. Shade Responsive Genes, Gene Components And Products
- V.P. Sulfur Responsive Genes, Gene Components And Products
- V.Q. Zinc Responsive Genes, Gene Components And Products
- VI. Utilities of Particular Interest
- 25 VII. Enhanced Foods
- VIII. Use Of Novel Genes To Facilitate Exploitation Of Plants As Factories For The Synthesis  
Of Valuable Molecules
- 30 IX. Promoters As Sentinels

## Description Of The Tables Containing The Data And Analytical Results

### How To Make Different Embodiments Of The Invention

- I. Polynucleotides And Methods Of Use Thereof
  - I.A. Probes, Primers And Substrates
  - I.B. Methods Of Detection And Isolation
    - I.B.1. Hybridization
    - I.B.2. Methods Of Mapping
    - I.B.3. Southern Blotting
    - I.B.4. Isolating Cdna From Related Organisms
    - I.B.5. Isolating And/Or Identifying Orthologous Genes
  - I.C. Methods Of Inhibiting Gene Expression
    - I.C.1. Antisense
    - I.C.2. Ribozyme Constructs
    - I.C.3. Chimeraplasts
    - I.C.4. Co-Suppression
    - I.C.5. Transcriptional Silencing
    - I.C.6. Other Methods To Inhibit Gene Expression
  - I.D. Methods Of Functional Analysis
  - I.E. Promoter Sequences And Their Use
  - I.F. Utrs And/Or Intron Sequences And Their Use
  - I.G. Coding Sequences And Their Use.
- II. Polypeptides And Proteins And Methods Of Use Thereof, Such As
  - II.A. Native Polypeptides And Proteins
    - II.A.1. Antibodies
    - II.A.2. In Vitro Applications
  - II.B. Polypeptide Variants, Fragments And Fusions
    - II.B.1. Variants

II.B.2. Fragments

II.B.3. Fusions

III. Methods Of Modulating Polypeptide Production, Such As

III.A. Suppression

III.A.1. Antisense

III.A.2. Ribozymes

III.A.3. Co-Suppression

III.A.4. Insertion Of Sequences Into The Gene To Be Modulated

III.A.5. Promoter Modulation

III.A.6. Expression Of Genes Containing Dominant-Negative Mutations

III.B. Enhanced Expression

III.B.1. Insertion Of An Exogenous Gene

III.B.2. Promoter Modulation

IV. Gene Constructs And Vector Construction, Such As

IV.A. Coding Sequences

IV.B. Promoters

IV.C. Signal Peptides

V. The Invention Still Further Relates To

V.A. Transformation Techniques

**Definitions**

**Examples**

**Claims**

**Abstract**

**Schematic 5**

POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
OF MODULATING VARIOUS RESPONSES

5 This application claims priority under 35 USC §119(e), §119(a-d) and §120 of the  
following applications:

Country	Attorney Docket No.	Application No.	Filed
United States	2750-1156P	60/228,279	8/25/00
United States	2750-2137P	Client No. 92004.005	8/25/00
United States	2750-2136P		8/25/00
United States	2750-2135P	60/228,246	8/25/00
United States	2750-1155P	60/228,224	8/25/00

This application contains a CDR, the entire contents of which are hereby incorporated by  
reference. The CDR contains the following files:

File Name	Date of Creation	File Size
010809 Protein Domain Table	August 09, 2001	2.62MB
2750-1481P AFLP_Diff Table 1	August 23, 2001	14.7KB
2750-1481P AFLP_Diff Table 2	August 23, 2001	8.73KB
2750-1481P AFLP_Diff Table 3	August 23, 2001	4.09KB
2750-1481P AFLP_Diff Table 4	August 23, 2001	7.85KB
2750-1481P AFLP_Diff Table 5	August 23, 2001	27.1KB
2750-1481P AFLP_Diff Table 6	August 23, 2001	24.3KB
2750-1481P AFLP_Diff Table 7	August 23, 2001	18.7KB
2750-1481P AFLP_Diff Table 8	August 23, 2001	36.3KB
2750-1481P AFLP_Diff Table 9	August 23, 2001	16.9KB
2750-1481P AFLP_Diff Table 10	August 23, 2001	33.2KB
2750-1481P Nitrogen Reference Table 1-01	August 23, 2001	4.42MB
2750-1481P Nitrogen Reference Table 1-02	August 23, 2001	2.81MB
2750-1481P Nitrogen Reference Table 1-03	August 23, 2001	43.9KB
2750-1481P Nitrogen Reference Table 1-04	August 23, 2001	27.0KB
2750-1481P GA Reference Table 1-05	August 23, 2001	1.36MB
2750-1481P GA Reference Table 1-06	August 23, 2001	2.7MB
2750-1481P GA Reference Table 1-07	August 23, 2001	15.8KB
2750-1481P GA Reference Table 1-08	August 23, 2001	22.7KB
2750-1481P SA Reference Table 1-09	August 23, 2001	7.88MB
2750-1481P SA Reference Table 1-10	August 23, 2001	7.28MB
2750-1481P SA Reference Table 1-11	August 23, 2001	103KB
2750-1481P SA Reference Table 1-12	August 23, 2001	106KB
2750-1481P Reference Table 1-13	August 23, 2001	6.15MB

File Name	Date of Creation	File Size
2750-1481P Reference Table 1-14	August 23, 2001	9.04MB
2750-1481P Reference Table 1-15	August 23, 2001	73.1KB
2750-1481P Reference Table 1-16	August 23, 2001	96.2KB
2750-1481P Reference Table 1-17	August 23, 2001	5.96MB
2750-1481P Reference Table 1-18	August 23, 2001	8.63MB
2750-1481P Reference Table 1-19	August 23, 2001	53.2KB
2750-1481P Reference Table 1-20	August 23, 2001	112KB
2750-1481P Nitrogen Sequence Table 2-01	August 23, 2001	6.30MB
2750-1481P Nitrogen Sequence Table 2-02	August 23, 2001	4.48MB
2750-1481P Nitrogen Sequence Table 2-03	August 23, 2001	404KB
2750-1481P Nitrogen Sequence Table 2-04	August 23, 2001	215KB
2750-1481P GA Sequence Table 2-05	August 23, 2001	12.12MB
2750-1481P GA Sequence Table 2-06	August 23, 2001	4.9MB
2750-1481P GA Sequence Table 2-07	August 23, 2001	142KB
2750-1481P GA Sequence Table 2-08	August 23, 2001	193KB
2750-1481P SA Sequence Table 2-09	August 23, 2001	12.6MB
2750-1481P SA Sequence Table 2-10	August 23, 2001	10.7MB
2750-1481P SA Sequence Table 2-11	August 23, 2001	857KB
2750-1481P SA Sequence Table 2-12	August 23, 2001	911KB
2750-1481P Sequence Table 2-13	August 23, 2001	10.6MB
2750-1481P Sequence Table 2-14	August 23, 2001	13.9MB
2750-1481P Sequence Table 2-15	August 23, 2001	840KB
2750-1481P Sequence Table 2-16	August 23, 2001	800KB
2750-1481P Sequence Table 2-17	August 23, 2001	12.2MB
2750-1481P Sequence Table 2-18	August 23, 2001	11.7MB
2750-1481P Sequence Table 2-19	August 23, 2001	658KB
2750-1481P Sequence Table 2-20	August 23, 2001	871KB
2750-1481P AFLP_Int Table 1	August 23, 2001	6.52MB
2750-1481P AFLP_Int Table 2	August 23, 2001	8.60MB

### FIELD OF THE INVENTION

The present invention relates to over 100,000 isolated polynucleotides that include a complete coding sequence, or a fragment thereof, that is expressed. These polynucleotides come from five plant species. In addition, the present invention relates to the polypeptide or protein corresponding to the coding sequence of these polynucleotides. The present invention also relates to isolated polynucleotides that represent regulatory regions of genes. The present invention also relates to isolated polynucleotides that represent untranslated regions of genes. The present invention further relates to the use of these isolated polynucleotides and polypeptides and proteins.

BACKGROUND AND SUMMARY OF THE INVENTION

There are more than 300,000 species of plants. They show a wide diversity of forms, ranging from delicate liverworts, adapted for life in a damp habitat, to cacti, capable of surviving in the desert. The plant kingdom includes herbaceous plants, such as corn, whose life cycle is measured in months, to the giant redwood tree, which can live for thousands of years. This diversity reflects the adaptations of plants to survive in a wide range of habitats. This is seen most clearly in the flowering plants (phylum Angiospermophyta), which are the most numerous, with over 250,000 species. They are also the most widespread, being found from the tropics to the arctic.

The process of plant breeding involving man's intervention in natural breeding and selection, is some 20,000 years old. It has produced remarkable advances in adapting existing species to serve new purposes. The world's economics was largely based on the successes of agriculture for most of these 20,000 years.

Plant breeding involves choosing parents, making crosses to allow recombination of gene (alleles) and searching for and selecting improved forms. Success depends on the genes/alleles available, the combinations required and the ability to create and find the correct combinations necessary to give the desired properties to the plant. Molecular genetics technologies are now capable of providing new genes, new alleles and the means of creating and selecting plants with the new, desired characteristics.

When the molecular and genetic basis for different plant characteristics are understood, a wide variety of polynucleotides, both endogenous polynucleotides and created variants, polypeptides, cells, and whole organisms, can be exploited to engineer old and new plant traits in a vast range of organisms including plants. These traits can range from the observable morphological characteristics, through adaptation to specific environments to biochemical composition and to molecules that the plants (organisms) exude. Such engineering can involve tailoring existing traits, such as increasing the production of taxol in yew trees, to combining traits from two different plants into a single organism, such as inserting the drought tolerance of a cactus into a corn plant. Molecular and genetic knowledge also allows the creation of new traits. For example, the production of chemicals and pharmaceuticals that are not native to particular species or the plant kingdom as a whole.

The application reports the inventions Applicants have discovered to build a foundation of scientific understanding of plant genomes to achieve these aims. These inventions include polynucleotide and polypeptide sequences, and data relating to where and when the genes are differentially expressed and phenotypic observations resulting from either aberrant gene  
5 activation or disruption. How these data are transformed into a scientific understanding of plant biology and the control of traits from a genetic perspective also is explained by the instant application. Applications of these discoveries to create new prototypes and products in the field of chemical, pharmaceutical, food, feed, and fiber production are described herein as well.

The achievements described in this application were possible because of the results from a cluster of technologies, a genomic engine, depicted below in Schematic 1, that allows information on each gene to be integrated to provide a more comprehensive understanding of gene structure and function and the deployment of genes and gene components to make new  
10 products.

a) The discoveries of the instant application

Applicants have isolated and identified over one hundred thousand genes, gene components and their products and thousands of promoters. The genes were isolated and/or characterized from arabidopsis, soybean, maize, wheat and rice. These species were selected because of their economic value and scientific importance and were deliberately chosen to  
15 include representatives of the evolutionary divergent dicotyledonous and monocotyledonous groups of the plant kingdom. The number of genes characterized in this application represents a large proportion of all the genes in these plant species.

The techniques used initially to isolate and characterize most of the genes, namely sequencing of full-length cDNAs, were deliberately chosen to provide information on complete coding  
20 sequences and on the complete sequences of their protein products.

Gene components and products the Applicants have identified include exons, introns, promoters, coding sequences, antisense sequences, terminators and other regulatory sequences. The exons are characterized by the proteins they encode and arabidopsis promoters are characterized by their position in the genomic DNA relative to where mRNA synthesis begins  
30 and in what cells and to what extent they promote mRNA synthesis.



Further exploitation of molecular genetics technologies have helped the Applicants to understand the functions and characteristics of each gene and their role in a plant. Three powerful molecular genetics approaches were used to this end:

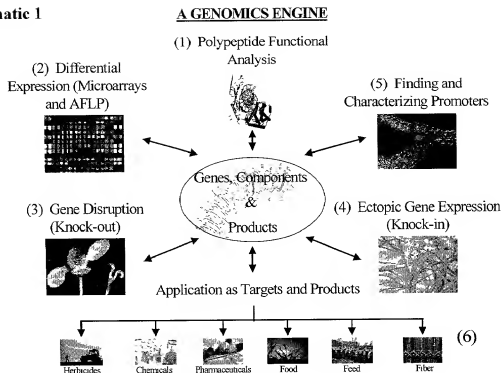
- (a) Analyses of the phenotypic changes when the particular gene sequence is interrupted or activated differentially; (arabidopsis)
- (b) Analyses of in what plant organs, to what extent, and in response to what environmental signals mRNA is synthesized from the gene; (arabidopsis and maize) and
- (c) Analysis of the gene sequence and its relatives. (all species)

These were conducted using the genomics engine depicted in Figure 1 that allows information on each gene to be integrated to provide a more comprehensive understanding of gene structure and function and linkage to potential products.

The species arabidopsis was used extensively in these studies for several reasons: (1) the complete genomic sequence, though poorly annotated in terms of gene recognition, was being produced and published by others and (2) genetic experiments to determine the role of the genes in planta are much quicker to complete.

The phenotypic tables, MA tables, AFLP data table and reference tables and sequence tables indicate the results of these analyses and thus the specific functions and characteristics that are ascribed to the genes and gene components and products.

**Schematic 1**



*Schematic 1. Gene sequences were determined and are depicted to occupy the center of the figure. Five different sorts of technologies were deployed in the Genomics engine to discover the functions of the genes. (1) Computer-based comparisons of protein structural features. (2) Studies to discover where and when each gene and groups of genes are active. (3) Discovery of the phenotypic consequences of inactivating each gene. (4) Elucidation of the phenotypic consequences of activating a gene in a new way. (5) Discovery of the sequence and activity of promoters of the genes. All this information leads to knowledge of how to use the genes, and gene components to create new products for industrial applications. (6)*

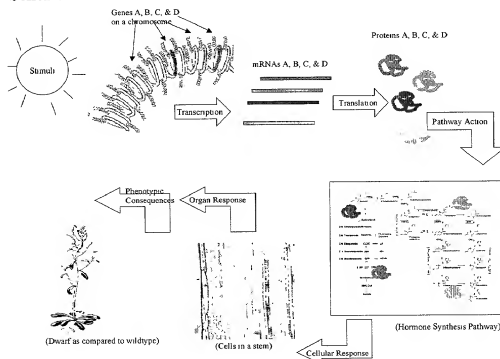
b) Integration of discoveries to provide scientific understanding

From the discoveries made, Applicants have deduced the biochemical activities, pathways, cellular roles, and developmental and physiological processes that can be modulated using these components. These are discussed and summarized in sections based on the gene functions characteristics from the analyses and role in determining phenotypes. These sections illustrate and emphasize that each gene, gene component or product influences biochemical activities, cells or organisms in complex ways, from which there can be many phenotypic consequences.

An illustration of how the discoveries on gene structure, function, expression and phenotypic observation can be integrated together to understand complex phenotypes is provided in Figure 2. This sort of understanding enables conclusions to be made as to how the genes, gene components and product are useful for changing the properties of plants and other organisms.

This example also illustrates how single gene changes in, for example, a metabolic pathway can cause gross phenotypic changes.

## Schematic 2



*Schematic 2. The figure illustrates how genes A, B, C and D are activated by internal stimuli and then their mRNA transcripts translated into proteins. These proteins are enzymes in three different but linked pathways. All three pathways are activated by the same stimuli. One of them, depicted by the green and light blue proteins determines the levels of a hormone in the shoot meristems causes cells to expand. This cell expansion leads to a longer stem and a taller plant. Genes A & C are therefore useful for controlling plant height and stem strength. The other two*

Furthermore, the development and properties of one part of plant can be interconnected with other parts. The dependence of shoot and leaf development on root cells is a classic example. Here, shoot growth and development require nutrients supplied from roots, so the protein complement of root cells can affect plant development, including flowers and seed production. Similarly, root development is dependent on the products of photosynthesis from

leaves. Therefore, proteins in leaves can influence root developmental physiology and biochemistry.

Thus, the Utility and Application sections describe both the functions and characteristics of the genes, gene components and products and also the multiplicity of biochemical activities, cellular functions, and the developmental and physiological processes influenced by them. The sections also describe examples of commercial products that can be realized from the inventions.

Analyses to Reveal Function and *in vivo* roles of single genes in one plant species.

The genomics engine has focused on individual genes to reveal the multiple functions or characteristics that are associated to each gene, gene components and products of the instant invention in the living plant. For example, the biochemical activity of a protein is deduced based on its similarity to a protein of known function. In this case, the protein may be ascribed with, for example, an oxidase activity. Where and when this same protein is active can be uncovered from differential expression experiments, which show that the mRNA encoding the protein is differentially expressed in response to drought and in seeds but not roots. The gene disruption experiments reveal that absence of the same protein causes embryo lethality.

Thus, this protein is characterized as a seed protein and drought-responsive oxidase that is critical for embryo viability.

Analyses to Reveal Function and roles of single genes in different species.

The genomics engine has also been used to extrapolate knowledge from one species to many plant species. For example, proteins from different species, capable of performing identical or similar functions, preserve many features of amino acid sequence and structure during evolution. Complete protein sequences have been compared and contrasted within and between species to determine the functionally vital domains and signatures characteristic of each of the proteins that is the subject of this application. Thus, functions and characteristics of arabidopsis proteins have been extrapolated to proteins containing similar domains and signatures of corn, soybean, rice and wheat and by implication to all other (plant) species.

Schematic 3 provides an example. Two proteins with related structures, one from corn, a monocot, and one from arabidopsis, a dicot, have been concluded to be orthologs. The known

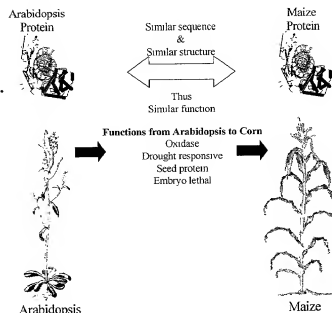
characteristics of the arabidopsis protein (seed protein, drought responsive oxidase) can then be

attributed to the  
protein.

CORN

**Integration of Data Across Species to Link Gene Products and Phenotypes**

**Schematic 3.**



Analyses over multiple experiments to reveal gene networks and links across species

The genomics engine can identify networks or pathways of genes concerned with the same process and hence linked to the same phenotype(s). Genes specifying functions of the same pathway or developmental environmental responses are frequently co-regulated i.e. they are regulated by mechanisms that result in coincident increases or decreases for all gene members in the group. The Applicants have divided the genes of arabidopsis and maize into such co-regulated groups on the basis of their expression patterns and the function of each group has been deduced. This process has provided considerable insight into the function and role of thousands of the plant genes in diverse species included in this application.

Applications of Applicant's discoveries

It will be appreciated while reading the sections that the different experimental molecular genetic approaches focused on different aspects of the pathway from gene and gene product through to the properties of tissues, organs and whole organisms growing in specific environments. For each endogenous gene, these pathways are delineated within the existing biology of the species. However, Applicants' inventions allow gene components or products to be mixed and matched to create new genes and placed in other cellular contexts and species, to exhibit new combinations of functions and characteristics not found in nature, or to enhance and modify existing ones. For instance, gene components can be used to achieve expression of a specific protein in a new cell type to introduce new biochemical activities, cellular attributes or developmental and physiological processes. Such cell-specific targeting can be achieved by combining polynucleotides encoding proteins with any one of a large array of promoters to facilitate synthesis of proteins in a selective set of plant cells. This emphasizes that each gene, component and protein can be used to cause multiple and different phenotypic effects depending on the biological context. The utilities are therefore not limited to the existing in vivo roles of the genes, gene components, and gene products.

While the genes, gene components and products disclosed herein can act alone, combinations are useful to modify or modulate different traits. Useful combinations include different polynucleotides and/or gene components or products that have (1) an effect in the same or similar developmental or biochemical pathways; (2) similar biological activities; (3) similar transcription profiles; or (4) similar physiological consequences.

Of particular interest are the transcription factors and key factors in regulatory transduction pathways, which are able to control entire pathways, segments of pathways or large groups of functionally related genes. Therefore, manipulation of such proteins, alone or in combination is especially useful for altering phenotypes or biochemical activities in plants. Because interactions exist between hormone, nutrition, and developmental pathways, combinations of genes and/or gene products from these pathways also are useful to produce more complex changes. In addition to using polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may exhibit different transcription profiles but which participate in common or overlapping pathways. Also, polynucleotides encoding selected enzymes can be combined in novel ways in a plant to create new metabolic pathways and hence new metabolic products.

The utilities of the various genes, gene components and products of the Application are described in the sections as follows:

I. Organ Affecting Genes, Gene Components, Products (Including Differentiation Function)

I.A. Root Genes, Gene Components And Products

I.A.1. Root Genes, Gene Components And Products

I.A.2. Root Hair Genes, Gene Components And Products

I.B. Leaf Genes, Gene Components And Products

I.B.1. Leaf Genes, Gene Components And Products

I.B.2. Trichome Genes And Gene Components

I.B.3. Chloroplast Genes And Gene Components

I.C. Reproduction Genes, Gene Components And Products

I.C.1. Reproduction Genes, Gene Components And Products

I.C.2. Ovule Genes, Gene Components And Products

I.C.3. Seed And Fruit Development Genes, Gene Components And Products

I.D. Development Genes, Gene Components And Products

I.D.1. Imbibition and Germination Responsive Genes, Gene Components And Products

I.D.2. Early Seedling Phase Genes, Gene Components And Products

I.D.3. Size and Stature Genes, Gene Components And Products

I.D.4. Shoot-Apical Meristem Genes, Gene Components And Products

I.D.5. Vegetative-Phase Specific Responsive Genes, Gene Components And Products

II. Hormones Responsive Genes, Gene Components And Products

II.A. Abscissic Acid Responsive Genes, Gene Components And Products

II.B. Auxin Responsive Genes, Gene Components And Products

II.C. Brassinosteroid Responsive Genes, Gene Components And Products

II.D. Cytokinin Responsive Genes, Gene Components And Products

II.E. Gibberellic Acid Responsive Genes, Gene Components And Products

5 III. Metabolism Affecting Genes, Gene Components And Products

III.A. Nitrogen Responsive Genes, Gene Components And Products

III.B. Circadian Rhythm Responsive Genes, Gene Components And Products

III.C. Blue Light (Phototropism) Responsive Genes, Gene Components And Products

III.D. Co2 Responsive Genes, Gene Components And Products

10 III.E. Mitochondria Electron Transport Genes, Gene Components And Products

III.F. Protein Degradation Genes, Gene Components And Products

III.G. Carotenogenesis Responsive Genes, Gene Components And Products

15 IV. Viability Genes, Gene Components And Products

IV.A. Viability Genes, Gene Components And Products

IV.B. Histone Deacetylase (Axel) Responsive Genes, Gene Components And Products

20 V. Stress Responsive Genes, Gene Components And Products

V.A. Cold Responsive Genes, Gene Components And Products

V.B. Heat Responsive Genes, Gene Components And Products

V.C. Drought Responsive Genes, Gene Components And Products

V.D. Wounding Responsive Genes, Gene Components And Products

V.E. Methyl Jasmonate Responsive Genes, Gene Components And Products

V.F. Reactive Oxygen Responsive Genes, Gene Components And H2O2 Products

25 V.G. Salicylic Acid Responsive Genes, Gene Components And Products

V.H. Nitric Oxide Responsive Genes, Gene Components And Products

V.I. Osmotic Stress Responsive Genes, Gene Components And Products

V.J. Aluminum Responsive Genes, Gene Components And Products

V.K. Cadmium Responsive Genes, Gene Components And Products

30 V.L. Disease Responsive Genes, Gene Components And Products

V.M. Defense Responsive Genes, Gene Components And Products



V.N. Iron Responsive Genes, Gene Components And Products  
V.O. Shade Responsive Genes, Gene Components And Products  
V.P. Sulfur Responsive Genes, Gene Components And Products  
V.Q. Zinc Responsive Genes, Gene Components And Products

5

VI. Enhanced Foods

VII. Pharmaceutical Products

10

VIII. Precursors Of Industrial Scale Compounds

IX. Promoters As Sentinels

## SUMMARY OF THE INVENTION

5 The present invention comprises polynucleotides, such as complete cDNA sequences and/or sequences of genomic DNA encompassing complete genes, fragments of genes, and/or regulatory elements of genes and/or regions with other functions and/or intergenic regions, hereinafter collectively referred to as Sequence-Determined DNA Fragments (SDFs) or sometimes collectively referred to as "genes or gene components", or sometimes as "genes, gene components or products", from different plant species, particularly corn, wheat, soybean, rice and *Arabidopsis thaliana*, and other plants and or mutants, variants, fragments or fusions of said SDFs and polypeptides or proteins derived therefrom. In some instances, the SDFs span the entirety of a protein-coding segment. In some instances, the entirety of an mRNA is represented. Other objects of the invention that are also represented by SDFs of the invention are control sequences, such as, but not limited to, promoters. Complements of any sequence of the invention are also considered part of the invention.

10 Other objects of the invention are polynucleotides comprising exon sequences, polynucleotides comprising intron sequences, polynucleotides comprising introns together with exons, intron/exon junction sequences, 5' untranslated sequences, and 3' untranslated sequences of the SDFs of the present invention. Polynucleotides representing the joinder of any exons described herein, in any arrangement, for example, to produce a sequence encoding any desirable amino acid sequence are within the scope of the invention.

15 The present invention also resides in probes useful for isolating and identifying nucleic acids that hybridize to an SDF of the invention. The probes can be of any length, but more typically are 12-2000 nucleotides in length; more typically, 15 to 200 nucleotides long; even more typically, 18 to 100 nucleotides long.

20 Yet another object of the invention is a method of isolating and/or identifying nucleic acids using the following steps:

- 25 (a) contacting a probe of the instant invention with a polynucleotide sample under conditions that permit hybridization and formation of a polynucleotide duplex; and  
(b) detecting and/or isolating the duplex of step (a).

30 The conditions for hybridization can be from low to moderate to high stringency conditions. The sample can include a polynucleotide having a sequence unique in a plant genome. Probes and

methods of the invention are useful, for example, without limitation, for mapping of genetic traits and/or for positional cloning of a desired fragment of genomic DNA.

Probes and methods of the invention can also be used for detecting alternatively spliced messages within a species. Probes and methods of the invention can further be used to detect or isolate related genes in other plant species using genomic DNA (gDNA) and/or cDNA libraries. In some instances, especially when longer probes and low to moderate stringency hybridization conditions are used; the probe will hybridize to a plurality of cDNA and/or gDNA sequences of a plant. This approach is useful for isolating representatives of gene families which are identifiable by possession of a common functional domain in the gene product or which have common cis-acting regulatory sequences. This approach is also useful for identifying orthologous genes from other organisms.

The present invention also resides in constructs for modulating the expression of the genes comprised of all or a fragment of an SDF. The constructs comprise all or a fragment of the expressed SDF, or of a complementary sequence. Examples of constructs include ribozymes comprising RNA encoded by an SDF or by a sequence complementary thereto, antisense constructs, constructs comprising coding regions or parts thereof, constructs comprising promoters, introns, untranslated regions, scaffold attachment regions, methylating regions, enhancing or reducing regions, DNA and chromatin conformation modifying sequences, etc. Such constructs can be constructed using viral, plasmid, bacterial artificial chromosomes (BACs), plasmid artificial chromosomes (PACs), autonomous plant plasmids, plant artificial chromosomes or other types of vectors and exist in the plant as autonomous replicating sequences or as DNA integrated into the genome. When inserted into a host cell the construct is, preferably, functionally integrated with, or operatively linked to, a heterologous polynucleotide. For instance, a coding region from an SDF might be operably linked to a promoter that is functional in a plant.

The present invention also resides in host cells, including bacterial or yeast cells or plant cells, and plants that harbor constructs such as described above. Another aspect of the invention relates to methods for modulating expression of specific genes in plants by expression of the coding sequence of the constructs, by regulation of expression of one or more endogenous genes in a plant or by suppression of expression of the polynucleotides of the invention in a plant. Methods of modulation of gene expression include without limitation (1) inserting into a host cell additional

copies of a polynucleotide comprising a coding sequence; (2) modulating an endogenous promoter in a host cell; (3) inserting antisense or ribozyme constructs into a host cell and (4) inserting into a host cell a polynucleotide comprising a sequence encoding a variant , fragment, or fusion of the native polypeptides of the instant invention.

## DETAILED DESCRIPTION OF THE INVENTION

### BRIEF DESCRIPTION OF THE INVENTIONS

As noted above, the Applicants have obtained and analyzed an extensive amount of information on a large number of genes by use of the Ceres Genomic Engine to determine. This information can be categorized into three basic types:

- I. Sequence Information for the Inventions
- II. Transcriptional Information for the Inventions
- III. Phenotypic Information for the Inventions

#### I. Sequence Information

##### Introduction to Sequence Information

To harness the potential of the plant genome, Applicants began by elucidating a large number gene sequences, including the sequences of gene components and products, and analyzing the data. The list of sequences and associated data are presented in the Reference and Sequence Tables of the present application (sometimes referred to as the "REF" and "SEQ" Tables). The Reference and Sequence tables include:

- cDNA sequence;
- coding sequence;
- 5' & 3' UTR;
- transcription start sites;
- exon and intron boundaries in genomic sequence; and
- protein sequence.

The Reference and Sequence Tables also include computer-based, comparative analyses between the protein sequences of the invention and sequences with known function. Proteins with similar sequences typically exhibit similar biochemical activities. The Reference table notes:

- sequences of known function that are similar to the Applicants' proteins; and
- biochemical activity that is associated with Applicants' proteins.

Also, by analyzing the protein sequences, Applicants were able to group the protein sequences into groups, wherein all the sequences in the group contain a signature sequence. The

groups are presented in the Protein Group Table. The signature sequences are reported in the Protein Group Table. More detailed analyses of the signature sequences are shown in the Protein Group Matrix Table.

10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995  
1000

IA. IDENTIFICATION OF GENE COMPONENTS AND PRODUCTS

To generate the Reference and Sequence Tables, Applicants took a cDNA/coding sequence approach. That is, Applicants initiated their studies either by isolating cDNAs and determining their sequences experimentally, or by identifying the coding sequence from genomic sequence with the aid of predictive algorithms. The cDNA sequences and coding sequences also are referred to as "Maximum Length Sequences" in the Reference tables. The cDNA and coding sequences were given this designation to indicate these were the maximum length of coding sequences identified by Applicants.

Due to this cDNA/coding sequence focus of the present application, the Reference and Sequence Tables were organized around cDNA and coding sequences. Each of these Maximum Length Sequences was assigned a unique identifier: Ceres Sequence ID NO, which is reported in the Tables.

All data that relate to these Maximum Length Sequences are grouped together, including 5' & 3' UTRs; transcription start sites; exon and intron boundaries in genomic sequence; protein sequence, etc.

Below, a more detailed explanation of the organization of the Reference and Sequence Tables and how the data in the tables were generated is provided.

a. cDNA

Applicants have ascertained the sequences of mRNAs from different organisms by reverse transcription of mRNA to DNA, which was cloned and then sequenced. These complementary DNA or cDNA sequences also are referred to as Maximum Length Sequences in the Reference Tables, which contain details on each of the sequences in the Sequence Tables.

Each sequence was assigned a Pat. Appln. Sequence ID NO: and an internal Ceres Sequence ID NO: as reported in the Reference Table, the section labeled "(Ac) cDNA Sequence." An example is shown below:

Max Len. Seq. :

(Ac) cDNA Sequence

- Pat. Appln. Sequence ID NO: 174538

- Ceres Sequence ID NO: 5673127

Both numbers are included in the Sequence Table to aid in tracking of information, as shown below:

<210> 174538 (Pat. Appln. Sequence ID NO:)

<211> 1846

<212> DNA (genomic)

<213> *Arabidopsis thaliana*

<220>

<221> misc\_feature

<222> (1)..(1846)

<223> Ceres Seq. ID no. 5673127

<220>

<221> misc\_feature

<222> ()..()

<223> n is a, c, t, g, unknown, or other

<400> 174538

acaagaacaa caaacacagag gaagaagaag aagaagatga agcttctggc tctgtttcca 60

tttctagcga tctgtgatcca actcagctgt... etc.

The Sequence and Reference Tables are divided into sections by organism: *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Triticum aestivum*; and *Oryza sativa*.

b. Coding Sequence

The coding sequence portion of the cDNA was identified by using computer-based

algorithms and comparative biology. The sequence of each coding sequence of the cDNA is reported in the "PolyP Sequence" section of the Reference Tables, which are also divided into sections by organism. An example shown below for the peptides that relate to the cDNA sequence above

PolyP Sequence

- Pat. Appln. Sequence ID NO 174539
- Ceres Sequence ID NO 5673128
- Loc. Sequence ID NO 174538: @ 1 nt.
- Loc. Sig. P. Sequence ID NO 174539: @ 37 aa.

The polypeptide sequence can be found in the Sequence Tables by either the Pat. Appln. Sequence ID NO or by the Ceres Sequence ID NO: as shown below:

<210> 174539 (Pat. Appln. Sequence ID NO)

<211> 443

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> peptide

<222> (1)..(443)

<223> Ceres Seq. ID no. 5673128

<220>

<221> misc\_feature

<222> ()..()

<223> xaa is any aa, unknown or other

<400> 174539

Thr Arg Thr Thr Lys Gln Arg Lys Lys Lys Lys Lys Met Lys Leu Leu



1                    5                    10                    15

Ala Leu Phe Pro Phe Leu Ala Ile... etc.

25

5

The PolyP section also indicates where the coding region begins in the Maximum Length Sequence. More than one coding region may be indicated for a single polypeptide due to multiple potential translation start codons. Coding sequences were identified also by analyzing genomic sequence by predictive algorithms, without the actual cloning of a cDNA molecule from a mRNA. By default, the cDNA sequence was considered the same as the coding sequence, when Maximum Length Sequence was a spliced together from a genomic annotation.

10

c. 5' and 3' UTR

The 5' UTR can be identified as any sequence 5' of the initiating codon of the coding sequence in the cDNA sequence. Similarly, the 3' UTR is any sequence 3' of the terminating codon of the coding sequence.

15

d. Transcription Start Sites

Applicants cloned a number of cDNAs that encompassed the same coding sequence but comprised 5' UTRs of different lengths. These different lengths revealed the multiple transcription start sites of the gene that corresponded to the cDNA. These multiple transcription start sites are reported in the "Sequence # w. TSS" section" of the Reference Tables.

20

e. Exons & Introns

Alignment of the cDNA sequences and coding portions to genomic sequence permitted Applicants to pinpoint the exon/intron boundaries. These boundaries are identified in the Reference Table under the "Pub gDNA" section. That section reports the gi number of the public BAC sequence that contains the introns and exons of interest.

25

Max Len. Seq. :

Pub gDNA:

gi No: 1000000005

30

Gen. seq. in cDNA:

115777 ... 115448 by Method #1  
115105 ... 114911 by Method #1  
114822 ... 114700 by Method #1  
114588 ... 114386 by Method #1  
114295 ... 113851 by Method #1  
115777 ... 115448 by Method #2  
115105 ... 114911 by Method #2  
114822 ... 114700 by Method #2  
114588 ... 114386 by Method #2  
114295 ... 113851 by Method #2  
115813 ... 115448 by Method #3  
115105 ... 114911 by Method #3  
114822 ... 114700 by Method #3  
114588 ... 114386 by Method #3  
114295 ... 113337 by Method #3

(Ac) cDNA Sequence

All the gi numbers were assigned by Genbank to track the public genomic sequences except:

gi 1000000001

gi 1000000002

gi 1000000003

gi 1000000004; and

gi 1000000005.

These gi numbers were assigned by Applicants to the five *Arabidopsis* chromosome sequences that were published by the Institute of Genome Research (TIGR). Gi 1000000001 corresponds to chromosome 1, Gi 1000000002 to chromosome 2, etc.

The method of annotation is indicated as well as any similar public annotations.

f. Promoters & Terminators

Promoter sequences are 5' of the translational start site in a gene; more typically, 5' of the transcriptional start site or sites. Terminator sequences are 3' of the translational terminator codon; more typically, 3' of the end of the 3' UTR.

For even more specifics of the Reference and Sequence Tables, see the section below titled "Brief Description of the Tables."

## II. TRANSCRIPTIONAL (DIFFERENTIAL EXPRESSION) INFORMATION

### Introduction to Differential Expression Data & Analyses

A major way that a cell controls its response to internal or external stimuli is by regulating the rate of transcription of specific genes. For example, the differentiation of cells during organogenesis into forms characteristic of the organ is associated with the selective activation and repression of large numbers of genes. Thus, specific organs, tissues and cells are functionally distinct due to the different populations of mRNAs and protein products they possess. Internal signals program the selective activation and repression programs. For example, internally synthesized hormones produce such signals. The level of hormone can be raised by increasing the level of transcription of genes encoding proteins concerned with hormone synthesis.

To measure how a cell reacts to internal and/or external stimuli, individual mRNA levels can be measured and used as an indicator for the extent of transcription of the gene. Cells can be exposed to a stimulus, and mRNA can be isolated and assayed at different time points after stimulation. The mRNA from the stimulated cells can be compared to control cells that were not stimulated. The mRNA levels of particular Maximum Length Sequences that are higher in the stimulated cell versus the control indicate a stimulus-specific response of the cell. The same is true of mRNA levels that are lower in stimulated cells versus the control condition.

Similar studies can be performed with cells taken from an organism with a defined mutation in their genome as compared with cells without the mutation. Altered mRNA levels in the mutated cells indicate how the mutation causes transcriptional changes. These transcriptional changes are associated with the phenotype that the mutated cells exhibit that is different from the phenotype exhibited by the control cells.

Applicants have utilized microarray and AFLP techniques to measure the levels of mRNAs in cells from mutant plants, stimulated plants, and/or selected from specific organs. The differential

expression of various genes in the samples versus controls are listed in the MA\_diff and/or AFLP\_diff and AFLP\_diff Tables.

Applicants have analyzed the differential data to identify genes whose mRNA transcription levels are positively correlated. From these analyses, Applicants were able to group different genes together whose transcription patterns are correlated. The results of the analyses are reported in the MA\_clust Tables.

a. Experimental Detail

A microarray is a small solid support, usually the size of a microscope slide, onto which a number of polynucleotides have been spotted onto or synthesized in distinct positions on the slide (also referred to as a chip). Typically, the polynucleotides are spotted in a grid formation. The polynucleotides can either be Maximum Length Sequences or shorter synthetic oligonucleotides, whose sequence is complementary to specific Maximum Length Sequence entities. A typical chip format is as follows:

Oligo #1	Oligo #2	Oligo #3
Oligo #4	Oligo #5	Oligo #6
Oligo #7	Oligo #8	Oligo #9

For Applicants' experiments, samples were hybridized to the chips using the "two-color" microarray procedure. A fluorescent dye was used to label cDNA reverse-transcribed from mRNA isolated from cells that had been stimulated, mutated, or collected from a specific organ or developmental stage. A second fluorescent dye of another color was used to label cDNA prepared from control cells.

The two differentially-labeled cDNAs were mixed together. Microarray chips were incubated with this mixture. For Applicants' experiments the two dyes that are used are Cy3, which fluoresces in the red color range, and Cy5, which fluoresces in the green/blue color range. Thus, if:

- cDNA#1 binds to Oligo #1;
- cDNA#1 from the sample is labeled red;
- cDNA#1 from the control is labeled green, and
- cDNA#1 is in both the sample and control,

then cDNA#1 from both the sample and control will bind to Oligo#1 on the chip. If the sample has 10 times more cDNA#1 than the control, then 10 times more of the cDNA#1 would be hybridized to Oligo#1. Thus, the spot on the chip with Oligo#1 spot would look red.

Oligo #1	Oligo #2	Oligo #3
Oligo #4	Oligo #5	Oligo #6
Oligo #7	Oligo #8	Oligo #9

If the situation were reversed, the spot would appear green. If the sample has approximately the same amount of cDNA#1 as the control, then the Oligo#1 spot on the chip would look yellow.

These color differentials are measured quantitatively and used to deduce the relative concentration of mRNAs from individual genes in particular samples.

b. MA\_diff and/or AFLP\_diff Data Table

To generate data, Applicants labeled and hybridized the sample and control mRNA in duplicate experiments. One chip was exposed to a mixture of cDNAs from both a sample and control, where the sample cDNA was labeled with Cy3, and the control was labeled with Cy5 dye. For the second labeling and chip hybridization experiments, the fluorescent labels were reversed; that is, the Cy5 dye for the sample, and the Cy3 dye for the control.

Whether Cy5 or Cy3 was used to label the sample, the fluorescence produced by the sample was divided by the fluorescence of the control. A cDNA was determined to be differentially expressed in response to the stimulus in question if a statistically-significant difference in the sample versus the control was measured by both chip hybridization experiments.

The MA\_diff and/or AFLP\_diff tables show which cDNA were significantly up-regulated as designated by a "+" and which were significantly down-regulated as designated by a "-" for each pair of chips using the same sample and control.

### III. PHENOTYPIC INFORMATION

One means of determining the phenotypic effect of a gene is either to insert extra active

copies of the gene or coding sequence, or to disrupt an existing copy of the gene in a cell or organism and measure the effects of the genetic change on one or more phenotypic characters or traits. "Knock-in" is used herein to refer to insertion of additional active copies of a gene or coding sequence. "Knock-out" refers to a plant where an endogenous gene(s) is disrupted. Applicants have used both methods of addition or disruption to determine the phenotypic effects of gene or gene components or products, and have thereby discovered the function of the genes and their utilities.

1. Knock-in results

The coding sequence of a desired protein can be functionally linked to a heterologous promoter to facilitate expression. Here, Applicants have operably linked a number of coding sequences to either one of the promoters listed below:

<u>GFP Pattern</u>	<u>Specific Promoter activity</u>	<u>Plant Line Descriptor</u>
Root epidermis / mostly toward the lower region of root (more intense than CS9094)	Specific to the root basal region.	Root basal
Root-endodermis/cortex (initials sharp); shoot-mesophyll of one leaf, sharp guard cell marking. New leaf petioles near tip of primary inflorescence; floral stems; in flowers at base of sepal, anther stems, and pistil	Specific to the root endodermis-cortex region, leaf petiole, and flowers.	Root/Petiole/Flowers
Broad root exp. (some dermal, some cortical, some vascular); shoot apex. Faintly in petiole; stem	Specific to root and stem.	Root/Stem1
High expression in stem, excluded from 1st true leaves/High in root. Faint expression in stem	Specific to stem and root.	Root/Stem2
Shoot meristem / whole root region; little bit on cotyledons. Base of leaves(axillary meristem?); base of sepals; inflorescence meristem; small amount in unfertilized pistil.	Specific to roots, shoot meristem, base of leaves and flowers.	Root/Stem/Leaves/Flowers
root tip vascular initials; vascular system throughout plant; Bud petal vasculature and pistil septum; Flower petal vasculature; Flower pistil septum; Pre fertilization ovules; Post fertilization ovule at chalazal end; Developing seed (young, maturing siliques); Seed coat and young embryos. GFP not observed in mature embryos.	Specific to vascular systems.	Vascular/Ovule/Young Seed/Embryo

<b>GFP Pattern</b>	<b>Specific Promoter activity</b>	<b>Plant Line Descriptor</b>
Flower, sepal / vascular tissue of root, stem, and cotyledons. Stems of new flowers; vasculature or petals, anthers, sepals, and pistil/silique; Vasculature throughout seedling: root, hypocotyl, petioles, stem, cotyledons, first true leaves; Rosette vasculature; Cauline leaf vasculature; Bud pedicel vasculature; Flower vasculature: (sepals, petals, filaments, pistil); Bud vasculature (sepal, petal, filament, pistil); Funiculus in both flower and bud; Some possible seed coat expression; Silique funiculus; Very faint fluorescence in mature embryo (auto fluorescence perhaps);	Specific to flowers, seed and vasculature.	Flowers/Seed/Vasculature/Embryo
Root expression - primarily in cortex (upper region of the root). No shoot expression	Specific to root.	Roots2
Root expression - less intense in whole root of young seedling. Shoot apical meristem; organ primordia in SAM region.	Specific to root and shoot apical meristem.	Root/SAM
Root epidermis/tip; shoot epidermis/vascular; leaf epidermis; expression in developing seed/ovule - mature embryo; Primary and lateral root cortex; Very strong in root cap; Base of flower bud and epidermis of carpels; Base of flower, epidermis of filaments, epidermis of carpels; Trichomes; Weak (hardly detectable) gfp expression in vasculature throughout seedling; Strong expression in trichomes; POST- fertilization SEED only; GFP strength increases as silique matures; Weak at suspensor end of the embryo; GFP observed in seed coat; Root and post fertilization seed specific gfp expression; Expression in seed coat.	Specific to seed and to epidermal layers of roots, shoots and leaves.	Seed/Epidermis/Ovary/Fruit
Young root dermis; dermal/cortical?/vascular in older root; general (epidermal?) shoot expression; ovules. some in sepals; vasculature of stem	Specific to roots, shoots, and ovules.	Roots/Shoots/Ovule
Vascular tissue of root; Meristem tissues: axillary meristems, floral meristems, base of flowers/sepals; Weak expression in hypocotyl, petiole and cotyledon vasculature..	Specific to root structural leaf vascular region and to floral buds and axillary meristem	Vasculature/Meristem

The chimeric constructs were transformed into *Arabidopsis thaliana*. The resulting transformed lines were screened to determine what phenotypes were changed due to introduced transgene. The phenotype changes, relative to the control, are reported in the Knock-in tables.

## 2. Knock-out Results

Knock-out plants in *Arabidopsis thaliana* were created by inserting a polynucleotide tag into the genome. The location of the tag was identified using primers to the tag sequence and isolation of the plant genomic sequence that flanks the tag using a variation of the polymerase chain reaction. The plants were generated using the procedure described in Feldmann et al., (1987) Molec. Gen. Genet. 208: 1-9; Feldmann (1991) Plant Journal, 1:71-83 and Forsthoefel et al., (1992) Aust. J. Plant Physiol. 19:353-366.. On average, the population of plants that was screened had ~1.5 to 2 tags. Generally, the number of tags ranged from 1 to greater than 5.

The polynucleotide tags were classified as either incorporated within a gene, or between two genes. The data in the Knock-out Table indicates which plants have a tag(s) causing a disruption in a gene, or a disruption between genes.

### a. Disruption in a Gene

For the sake of this analysis, the tag was considered to be causing a disruption in a gene when the tag was located:

- 1) less than 501 upstream of the transcriptional start site;
- 2) less than 701 upstream of the translational initiation codon;
- 3) between the translational initiation and termination codons of the gene,
- 4) less than 301 downstream of the translational stop codon; or
- 5) less than 151 downstream of a transcriptional termination site.

By this definition, a tag can be inserted in two genes. For example, if two genes have only 700 nucleotides between the translational termination codon of one gene and the



translational initiation codon of the other gene, the tag can be inserted into the terminator of one gene and the promoter of the other gene according to the definition above.

Genomic annotations by the method OCKHAM-OCDNA identify the transcriptional start and stop site of a gene.

5

b. Disruption between Genes

When a tag causes a disruption between two genes, either or both genes can be affected. Typically, a tag can affect a gene if it disrupts the genome at a location 3000 nt downstream to the start codon of a gene. More typically, insertions found 1000- 2000 nt upstream (5'), or 750- 1000 nt downstream (3') could be expected to disrupt expression.

10

c. More Than One Insert

A plant can have multiple tags. If a mutant phenotype is observed, then it can be attributed to any one or all of the tags.

15

HOW THE INVENTIONS REVEAL HOW GENES, GENE COMPONENTS AND PRODUCTS  
FUNCTION

The different experimental molecular genetic approaches focused on different aspects of genes, gene components, and gene products of the inventions. The variety of the data demonstrates the multiple functions and characteristics of single genes, gene components, and products. The data also explain the pathways and networks in which individual genes and products participate and interact. As a result, the circumstances or conditions are now known when these genes and networks are active. These new understandings of biology are relevant for many plant species. The following section describes the process by which Applicants analyzed the inventions generated by the Ceres Genomic Engine:

TABLE OF CONTENTS

- I. Experimental Results Reveal Many Facets Of A Single Gene
  - I.A. Intro: Functions And Characteristics Of Gene Components and Products Are Identified
    - I.A.1. Functions Of Coding Sequences Revealed By The Ceres Genomic Engine
      - A.1.a. Sequence Similarity To Proteins Of Known Function Can Be Used To Associate Biochemical Activities And Molecular Interaction To The Proteins Of The Invention
        - 1. Presence Of Amino Acid Motifs Indicates Biological Function
        - 2. Related Amino Acid Sequences Share Similar Biological Functions
      - A.1.b. Differential Expression Results Explain In Which Cellular Responses The Proteins Of The Invention Are Involved
        - 1. The Transcript Levels Of A Protein Over Time In Response To A Stimuli Are Revealed By Transcriptional Analyses Over Many Experiments.
        - 2. The Transcript Levels Of A Protein Over Different Developmental Stages Can Be Identified By Transcriptional Analyses Over Many Experiments

3. Proteins That Are Common In A Number Of Similar Responses Can Be Identified By Transcriptional Analyses Over A Number Of Experiments
4. Proteins That Are Common To Disparate Responses Can Be Identified By Transcriptional Analyses Over A Number Of Experiments

A.1.c. Observations Of Phenotypic Changes Show What Physiological Consequences Applicants' Proteins Can Produce

I.A.2. Specificity Of Regulatory Sequence Are Uncovered By The Ceres Genomic Engine

A.2.a. Differential Expression Results Explain Which External Or Internal Stimuli Trigger The Regulatory Sequences

II. Experimental Results Also Reveal Pathways Or Networks Of Genes

II.A. Genes Whose Transcription Are Well Coordinated Generally Act Together To Produce Proteins That Participate In The Same Pathway Or Network

II.A.1. Calculating The Correlation Coefficient Between Pairs Of Genes Based On The Differential Expression Data

II.A.2. The Complete Linkage Analyses of Differential Identity Genes With Similar Pattern Of Transcription

II.A.3. The Nearest Neighbor Analyses of Differential Group Genes With Correlated But Dissimilar Transcription Profiles

III. Experimental Results Reveal The Function And Characteristic Genes, Pathways And Networks

III.A. Linking Biochemical Or Metabolic Activities Of One Protein In A Cluster To The Other Proteins In The Same Microarray Cluster

III.B. Using Differential Expression Data To Determine When The Genes And Pathways Are Active

III.C. Using Phenotype Data To Determine When Genes And Pathways Are Active

IV. Experimental Results Provide An Understanding Of Genes, Pathways and Networks In Many Plant Species

IV.A. Intro: Using Protein Sequence Similarity To Assign Function From Arabidopsis Proteins To Proteins Of Other Species

IV.A.1. Conserved Sequences Between Proteins Of Different Species Give Rise To A Signature Sequence

IV.A.2. Linking Signature Sequences To Conservation Of Structural Characteristics

IV.A.3. Linking Signature Sequences To Conservation Of Biochemical Activities And Molecular Interactions

IV.A.4. Linking Signature Sequences To Conservation Of Cellular Responses Or Activities

IV.A.5. Linking Signature Sequences To Conservation Of Phenotypic Effects

I. EXPERIMENTAL RESULTS REVEAL MANY FACETS OF A SINGLE GENE

I.A. INTRO: FUNCTIONS AND CHARACTERISTICS OF GENE COMPONENTS AND PRODUCTS ARE IDENTIFIED

The experimental results are used to dissect the function of individual components and products of the genes. For example, the biochemical activity of the encoded protein could be surmised from sequence analyses, and promoter specificity could be identified through transcriptional analyses. Generally, the data presented herein can be used to functionally annotate either the protein sequence and/or the regulatory sequence that control transcription and translation.

I.A.1. FUNCTIONS OF CODING SEQUENCES REVEALED BY THE CERES GENOMIC ENGINE

A.1.a. SEQUENCE SIMILARITY TO PROTEINS OF KNOWN FUNCTION CAN BE USED TO ASSOCIATE BIOCHEMICAL

ACTIVITIES AND MOLECULAR INTERACTION TO THE  
PROTEINS OF THE INVENTION

The protein sequences of the invention were analyzed to determine if they shared any sequence characteristics with proteins of known activity. Proteins can be grouped together based on sequence similarity, either localized or throughout the length of the proteins. Typically, such groups of proteins exhibit common biochemical activities or interact with similar molecules.

1. PRESENCE OF AMINO ACID MOTIFS INDICATES  
BIOLOGICAL FUNCTION

Localized protein sequence similarity, also referred to as amino acid motifs, have been attributed to enzyme or protein functions. A library of motifs, important for function, have been documented in PROSITE, a public database available at <http://www.expasy.ch/prosite/>. This library includes descriptions of the motifs and their functions. The zinc finger motif is one such entry in PROSITE, which reports that the zinc finger domain of DNA-binding proteins is typically defined by a 25-30 amino acid motif containing specific cysteine or histidine residues that are involved in the tetrahedral coordination of a zinc ion. Any protein comprising a sequence similar to the zinc finger amino acid motif will have similar functional activity (specific binding of DNA).

Protein sequences of the invention have been compared to a library of amino acid motifs in the pFAM database, which is linked to the PROSITE database. If any of Applicants' protein sequences exhibit similarity to these amino acid motifs or domains, the Reference Table notes the name and location of the motif in the "Pred. PP Nom. & Annot" section of the Reference tables. A description of any biochemical activities that are associated to these domains, and therefore associated with Applicants' proteins, is included in the Protein Domain table.

For example, polypeptide, CERES Sequence ID NO: 1545823 is associated with zinc finger motif as follows in the Reference Table:

(C) Pred. PP Nom. & Annot.

- Zinc finger, C3HC4 type (RING finger)
- Loc. Sequence ID NO 133059: 58 -> 106 aa.

2. RELATED AMINO ACID SEQUENCES SHARE SIMILAR

### BIOLOGICAL FUNCTIONS

It is apparent, when studying protein sequence families, that some regions have been better conserved than others during evolution. These regions are generally important for the function of a protein and/or for the maintenance of its three-dimensional structure.

5 The Reference Table reports in section "(Dp) Rel. AA Sequence" when a protein shares amino acid similarity with a protein of known activity. The section reports the gi number of the protein of known activity, a brief description of the activity, and the location where it shares sequence similarity to Applicants' polypeptide sequence.

Using this analysis, biochemical activity of the known protein is associated with Applicants' proteins. An example for the polypeptide described above is as follows:

#### (Dp) Rel. AA Sequence

- Align. NO 524716
- gi No 2502079
- Desp. : (AF022391) immediate early protein; ICP0 [Feline herpesvirus 1]
- % Idnt. : 33.7
- Align. Len.: 87
- Loc. Sequence ID NO 133059: 52 -> 137 aa.

### A.1.b. DIFFERENTIAL EXPRESSION RESULTS EXPLAIN IN WHICH CELLULAR RESPONSES THE PROTEINS OF THE INVENTION ARE INVOLVED

Differential expression has been studied in the present instance by both microarray and AFLP approaches. Differential expression results show when the coding sequence is transcribed, and therefore when the activity of the protein is deployed by the cell. Similar coding sequences can have very different physiological consequences because the sequences are expressed at different times or places, rather than because of any differences in protein activity. Therefore, modified levels (increased or decreased) of expression as compared to a control provide an indication of the function of a corresponding gene, gene components, and gene products.

These experiments can determine which are genes "over-expressed" under a given stimulus. Such over-expressed genes give rise to higher transcript levels in a plant or cell that is stimulated as

compared to the transcript levels of the same genes in a control organism or cell. Similarly, differential expression experiments can reveal "under-expressed" genes.

To increase the cellular response to a stimulus, additional copies of the coding sequences of a gene that is over-expressed are inserted into a cell. Increasing transcript levels of an over-expressed gene can either heighten or prolong the particular cellular response. A similar enhancement can occur when transcription of an under-expressed gene is inhibited. In contrast, the cellular response will be shortened or less severe when the over-expressed genes are inhibited or when expression of the under-expressed genes are increased.

In addition to analyzing the levels of transcription, the data were also analyzed to gain insight into the changes in transcription over time. That is, while the plants in the experiments were reacting to either an external or internal stimulus, a differential experiment takes a snapshot of the transcription levels in the cells at one specific time. However, a number of snap-shots can be taken at different time points during an external stimulus regime, or at different stages of development during an internal stimulus. These results show how the plant changes transcription levels over time, and therefore protein levels in response to specific stimuli to produce phenotypic changes. These results show that a protein can be implicated in a single, but more likely, in a number of cellular responses.

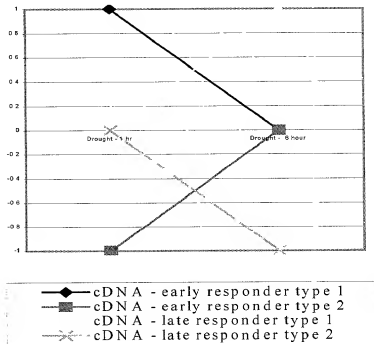
1. THE TRANSCRIPT LEVELS OF A PROTEIN OVER TIME  
IN RESPONSE TO A STIMULI ARE REVEALED BY  
TRANSCRIPTIONAL ANALYSES OVER MANY  
EXPERIMENTS

Applicants produced data from plants at different times after a specific stimulus. These results show whether the expression level of a gene spikes at a key moment during the cellular response, or whether the transcript level remains constant. Thus, coding sequences not only can be determined to be over- or under-expressed, but also can be classified by the initial timing and duration of differential expression. This understanding of timing can be used to increase or decrease any desired cellular response.

Generally, Applicants have assayed plants at 2 to 4 different time points after exposing the plants to the desired stimuli. From these experiments, "early" and "late" responders were identified. These labels are applied to either the regulatory sequences driving transcription of the gene as well

as to the protein encoded by the gene.

The following example illustrates how the genes, gene components and products were classified as either early or late responders following a specific. The mRNAs from plants exposed to drought conditions were isolated 1 hour and 6 hours after exposure to drought conditions. These mRNAs were tested utilizing microarray techniques. The graph below illuminates possible transcription profiles over the time course, plotting all the (+) data points as +1 and all the (-) data points as -1:



(The value for each time point was determined using a pair of microarray chips as described above.)

Data acquired from this type of time course experiment are useful to understand how one may increase or decrease the speed of the cellular response. Inserting into a cell extra copies of the coding sequence of early responders in order to over-express the specific gene can trigger a faster cellular response. Alternatively, coding sequences of late responders that are over-expressed can be placed under the control of promoters of early responders as another means to increase the cellular response.



Inserting anti-sense or sense mRNA suppression constructs of the early responders that are over-expressed can retard action of the late responders, thereby delaying the desired cellular response. In another embodiment, extra copies of the promoters of both early and late responders can be added to inhibit expression of both types of over-expressed genes.

The experiments described herein can be grouped together to determine the time course of the transcript levels of different coding sequences in response to different stimuli. Examples of different groups are as follows (the examples include the IDs for both corn and Arabidopsis experiments):

- NAA (EXPT IDs 108564, 108565, 108516, 108554)
- BA (EXPT IDs 108566, 108567, 108517)
- GA (EXPT IDs 108562, 108563, 108519, 108520, 108521, 108484, 108485, 108486)
- BR (EXPT IDs 108580, 108581, 108557, 108478, 108479, 108480, 108481)
- ABA (EXPT IDs 108560, 108561, 108513, 108597)
- Drought (EXPT IDs 108572, 108573, 108502, 108503, 108504, 108556, 108482, 108483, 108473, 108474, 108477)
- Cold (EXPT IDs 108578, 108579, 108533, 108534)
- Heat (EXPT IDs 108576, 108577, 108522, 108523)
- Osmotic stress (EXPT IDs 108570, 108571, 108541, 108542, 108553, 108539, 108540)
- Reactive Oxygen (EXPT IDs 108582, 108583, 108537, 108538, 108558)
- NO (EXPT IDs 108584, 108585, 108526, 108527, 108559)
- Wounding (EXPT IDs 108574, 108575, 108524, 108525)
- SA (EXPT IDs 108586, 108587, 108515, 108552, 108471, 108472, 108469, 108470, 107953, 107960, 108443, 108440, 108441, 108475, 108476)
- MeJA (EXPT IDs 108568, 108569)
- Finale (EXPT IDs 108467, 107871, 107876)
- Trimec (EXPT IDs 108466, 107886, 107891)
- Round-up (EXPT IDs 108465, 107896)
- Glean (EXPT IDs 108468, 107881)

## 2. THE TRANSCRIPT LEVELS OF A PROTEIN OVER

DIFFERENT DEVELOPMENTAL STAGES CAN BE  
IDENTIFIED BY TRANSCRIPTIONAL ANALYSES OVER  
MANY EXPERIMENTS

Differential expression data were produced for different development stages of various  
5 organs and tissues. Measurement of transcript levels can divulge whether specific genes give rise to  
spikes of transcription at specific times during development, or whether transcription levels remain  
constant. This understanding can be used to increase speed of development, or to arrest  
development at a specific stage.

10 Like the time-course experiments, the developmental stage data can classify genes as being  
transcribed at early or late stages of development. Generally, Applicants assayed different organs or  
tissues at 2-4 different stages.

15 Inhibiting under-expressed genes at either early or late stages can trigger faster development  
times. The overall development time also can be increased by this means to allow organs and tissue  
to grow to a larger size or to allow more organs or tissues to be produced. Alternatively, coding  
sequences of late stage genes that are under-expressed can be placed under the control of promoters  
of early stage genes to increase heighten development.

Inserting extra copies of the coding sequence early stage genes that are under-expressed can  
retard action of the late-stage genes and delay the desired development.

*cDNA\_ID 1595707*

*Cell Wall Synthesis  
Gene*



Fruit development of *Arabidopsis* is one example that can be studied. Siliques of varying sizes, which are representative of different stages, were assayed by microarray techniques. Specifically, mRNA was isolated from siliques between 0-5mm, between 5-10mm and > 10 mm in length. The graph below shows expression pattern of a cell wall synthesis gene, cDNAID 1595707, during fruit development:

The developmental course shows that the gene encoding a cell wall synthesis protein is up-regulated when the fruit is 0-5mm but returns to normal levels at 5-10mm and > 10mm. Increase of cell wall synthesis can lead to larger cells and/or greater number of cells. This type of increase can boost fruit yield. The coding sequence of the cell wall synthesis protein under the control of a strong early stage promoter would increase fruit size or number.

A pectinesterase gene was also differentially expressed during fruit development, cDNA ID 1396123. Pectinesterase catalyzes the hydrolysis of pectin into pectate and methanol. This biochemical activity plays an important role in cell wall metabolism during fruit ripening. To shorten the time for fruit ripening, extra copies of this gene with its endogenous promoter can be inserted into a desired plant. With its native promoter, the extra copies of the gene will be expressed at the normal time, to promote extra pectinesterase at the optimal stage of fruit development thereby shortening ripening time.

A number of Applicant's experiments can be grouped together to study changes of transcript levels over a number development stages. Below are examples of groups of experiments:

- Root, Root Tip, and rhl mutant (EXPT IDs 108594, 108433, 108599, 108434, 108439)
- Flowers Drought Exposed Flowers, SA Treated Flowers (EXPT IDs 108473, 108474, 108429, 108430, 108431, 108475, 108476, 108501)
- BR Shoot Apices, Leaves, Stm (EXPT IDs 108478, 108479, 108480, 108481, 108598, 108535, 108536, 108435)
- Leaf and Stm (EXPT IDs 108477, 108512, 108497, 108498, 108598, 108478, 108479, 108480, 108481, 108598, 108535, 108536, 108435)
- Imbibed & Germinating Seeds 1, 2, 3, And 4 Days (EXPT IDs 108461, 108462, 108463, 108464, 108528, 108529, 108530, 108531, 108545, 108546, 108547, 108518, 108529, 108543, 108544)
- Tissue Specific Expression (3 week rosette leaves, Tissue Specific Expression (3 week stems), Tissue Specific Expression (2 week roots) (EXPT IDs 108497, 108498, 108439)
- Tissue Specific Expression (3 week rosette leaves), Germinating Seeds (EXPT IDs 108497, 108461)
- Tissue Specific Expression (3 week rosette leaves, stm mutants, BR\_Shoot Apices Expt, root tips, Tissue Specific Expression (2 week roots) (EXPT IDs 108497, 108435, 108480, 108434, 108439)
- BR\_Shoot Apices Expt, root tips, Tissue Specific Expression (flower buds) (EXPT IDs 108480, 108434, 108431)
- Arab\_Ler-pi\_ovule\_1, ap2-floral buds, Tissue Specific Expression (flower buds), Tissue Specific Expression (<5 mm siliques) (EXPT IDs 108595, 108501, 108431, 108437)
- Tissue Specific Expression (2 week roots), rhl mutant2, BR\_Shoot Apices Expt, Trichome Inflorescences (EXPT IDs 108439, 108433, 108480, 108452)

3. PROTEINS THAT ARE COMMON IN A NUMBER OF  
SIMILAR RESPONSES CAN BE IDENTIFIED BY  
TRANSCRIPTIONAL ANALYSES OVER A NUMBER OF

## EXPERIMENTS

The differential expression experiments also reveal the genes, and therefore the coding sequence, that are common to a number of cellular responses. By identifying the genes that are differentially expressed in a number of similar responses, the genes at the nexus of a range of responses are discovered. For example, genes that are differentially expressed in all the stress responses are at the hub of many of the stress response pathways.

These types of nexus genes, proteins, and pathways are differentially expressed in many or majority of the responses or developmental conditions of interest. Typically, a nexus gene, protein, or pathway is differentially expressed in generally the same direction in many or majority of all the desired experiments. By doing so, the nexus gene can be responsible for triggering the same or similar set of pathways or networks for various cellular responses. This type of gene is useful in modulating pleiotropic effects or triggering or inhibiting a general class of responses.

When nexus genes are differentially expressed in a set of responses, but in different directions, these data indicate that a nexus gene is responsible for creating the specificity in a response by triggering the same pathway but to a different degree. Placing such nexus genes under a constitutive promoter to express the proteins at a more constant level can remove the fluctuations. For example, a plant that is better drought adapted, but not cold adapted can be modified to be tolerant to both conditions by placing under the control of a constitutive promoter a nexus gene that is up-regulated in drought but down regulated in cold.

Applicants' experiments can be grouped together to identify such nexus genes. Examples of these groups are as follows:

- Herbicide Response
  - Trimec, Finale, Glean, Round-up (EXPT IDs 108467, 107871, 107876, 108468, 107881, 108465, 107896, 108466, 107886, 107891)
- Stress Response
  - Drought, Cold, Heat, Osmotic Stress (EXPT IDs 108578, 108579, 108533, 108534, 108572, 108573, 108502, 108503, 108504, 108556, 108482, 108483, 108473, 108474, 108477, 108576, 108577, 108522, 108523, 108570, 108571, 108541, 108542, 108553, 108539, 108540)
  - Drought, Cold, Heat, PEG, Trimec, Finale, Glean, Round-up (EXPT IDs 108578,

108579, 108533, 108534, 108572, 108573, 108502, 108503, 108504, 108556,  
108482, 108483, 108473, 108474, 108477, 108576, 108577, 108522, 108523,  
108570, 108571, 108541, 108542, 108553, 108539, 108540)

- Wounding, SA, MeJA, Reactive Oxygen, NO (EXPT IDs 108568, 108569, 108555,  
108584, 108585, 108526, 108527, 108559, 108582, 108583, 108537, 108538,  
108558, 108586, 108587, 108515, 108552, 108471, 108472, 108469, 108470,  
107953, 107960, 108443, 108440, 108441, 108475, 108476, 108574, 108575,  
108524, 108525)

- Hormone Responses

- NAA, BA, BR, GA, TRIMEC (EXPT IDs 108566, 108567, 108517, 108580,  
108581, 108557, 108478, 108479, 108480, 108481, 108562, 108563, 108519,  
108520, 108521, 108484, 108485, 108486, 108564, 108565, 108516, 108554,  
108466, 107886, 107891)
- NAA, Trimec (EXPT IDs 108566, 108567, 108517, 108580, 108581, 108557,  
108478, 108479, 108480, 108481, 108562, 108563, 108519, 108520, 108521,  
108484, 108485, 108486, 108564, 108565, 108516, 108554, 108466, 107886,  
107891)

#### 4. PROTEINS THAT ARE COMMON TO DISPARATE RESPONSES CAN BE IDENTIFIED BY TRANSCRIPTIONAL ANALYSES OVER A NUMBER OF EXPERIMENTS

Phenotypes and traits result from complex interactions between cellular pathways and networks. Which pathways are linked by expression of common genes to specify particular traits can be discerned by identifying the genes that show differential expression of seemingly disparate responses or developmental stages. For example, hormone fluxes in a plant can direct cell patterning and organ development. Genes that are differentially expressed both in the hormone experiments and organ development experiments would be of particular interest to control plant development.

Examples Of Such Pathway Interactions Include:

- (i) The Interaction Between Stress Tolerance Pathways And Metabolism Pathways;

- (ii) Interaction Between Hormone Responses And Developmental Changes In The Plant;
- (iii) Interactions Between Nutrient Uptake And Developmental Changes;
- (iv) Mediation Of Stress Response By Hormone Responses; And
- (v) Interactions Between Stress Response And Development.

Applicant's experiments can be grouped together to identify proteins that participate in interacting pathways or networks. Specific groups of experiments include, for example:

(i) Stress & Metabolism

- Germinating Seeds (Day 1), Arab\_0.1uM\_Epi-Brass\_1, Arab\_NO3\_H-to-L\_1, Arab\_100uM\_GA3\_1 (EXPT IDs 108461, 108580, 108592, 108562)

(ii) Hormones & Development

- NAA, BA & Root Tips (EXPT IDs 108566, 108567, 108517, 108564, 108565, 108516, 108554, 108434, 108466, 107886, 107891)
- NAA, Roots & Root Tips (EXPT IDs 108564, 108565, 108516, 108554, 108599, 108434, 108439, 108466, 107886, 107891)
- NAA, BA, Roots And/Or Root Tips (EXPT IDs 108564, 108565, 108516, 108554, 108599, 108434, 108439, 108466, 107886, 107891, 108566, 108567, 108517)
- NAA, BA And Leaf (EXPT IDs 108566, 108567, 108517, 108518, 108529, 108512, 108497, 108498, 108598, 108564, 108565, 108516, 108554, 108466, 107886, 107891)
- NAA, BA, Leaves, Roots And/Or Root Tips (EXPT IDs 108566, 108567, 108517, 108518, 108529, 108512, 108497, 108498, 108598, 108564, 108565, 108516, 108554, 108466, 107886, 107891, 108599, 108434, 108439)
- ABA & Siliques (Of Any Size) (EXPT IDs 108560, 108561, 108513, 108597, 108436, 108437, 108438)
- GA, Imbibed & Germinating Seeds, ABA & Siliques (Of Any Size) (EXPT IDs 108560, 108561, 108513, 108597, 108562, 108563, 108519, 108520, 108521, 108484, 108485, 108486, 108461, 108462, 108463,

108464, 108528, 108529, 108530, 108531, 108545, 108546, 108547,  
108518, 108529, 108543, 108544, 108436, 108437, 108438)

- Tissue Specific Expression (3 week rosette leaves), Arab\_0.1uM\_Epi-Brass\_1, Arab\_100uM\_GA3\_1, Germinating Seeds (Day 1), (EXPT IDs 108461, 108497, 108580, 108562, 108461)

(iii) Nutrient Uptake And Development

- Any Or All Nitrogen Experiments With Siliques (Of Any Size) (EXPT IDs 108592, 108593, 108588, 108589, 108590, 108591, 108532, 108548, 108549, 108550, 108551, 108454, 108455, 108487, 108488, 108489, 108436, 108437, 108438)
- Any Or All Nitrogen Experiments With Roots Or Root Tips (EXPT IDs 108518, 108529, 108592, 108593, 108588, 108589, 108590, 108591, 108532, 108548, 108549, 108550, 108551, 108454, 108455, 108487, 108488, 108489, 108594, 108433, 108599, 108434, 108439)

(iv) Stress & Hormones

- ABA, Drought (EXPT IDs 108560, 108561, 108513, 108597, 108572, 108573, 108502, 108503, 108504, 108556, 108482, 108483, 108473, 108474, 108477)
- ABA, Drought, Cold, Heat, & Wounding (EXPT IDs 108560, 108561, 108513, 108597, 108578, 108579, 108533, 108534, 108572, 108573, 108502, 108503, 108504, 108556, 108482, 108483, 108473, 108474, 108477, 108576, 108577, 108522, 108523, 108574, 108575, 108524, 108525)
- Tissue Specific Expression (3 week rosette leaves), Arab\_100uM\_ABA\_1, Ws Arabidopsis Drought 2 days, Ws Arabidopsis Drought 4 days (EXPT IDs 108497, 108560, 108477, 108482)

(v) Stress & Hormones Stress & Hormones

- Nitrogen High transition to Low, Arab\_NO3\_H-to-L\_1, Tissue Specific Expression (<5mm siliques), Tissue Specific Expression (5-10mm siliques) (EXPT IDs 108455, 108592, 108437, 108436)



5

A.1.c. OBSERVATIONS OF PHENOTYPIC CHANGES SHOW WHAT  
PHYSIOLOGICAL CONSEQUENCES APPLICANTS'  
PROTEINS CAN PRODUCE

Another direct means of determining the physiological consequences of a protein is to make aberrant decreases or increases of its expression level in a cell. To this end, Applicants have produced plants where specific genes have been disrupted, or produced plants that include an extra expressed copy of the gene. The plants were then planted under various conditions to determine if any visible physiological changes are caused. These changes then are attributed to the changes in protein levels.

I.B. SPECIFICITY OF REGULATORY SEQUENCE ARE UNCOVERED BY THE  
CERES GENOMIC ENGINE

I.B.1. DIFFERENTIAL EXPRESSION RESULTS EXPLAIN WHICH  
EXTERNAL OR INTERNAL STIMULI TRIGGER THE REGULATORY  
SEQUENCES

Transcriptional studies can reveal the time and place that genes are expressed. Typically, regulatory sequences, such as promoters, introns, UTRs, etc., control when and in which cells transcription occurs. Differential studies can explain the temporal- and location-specific regulatory sequences that control transcription.

Using the experiments that are provided herein, one skilled in the art can choose a promoter or any other regulatory sequence that is capable of facilitating the desired pattern of transcription. For example, if a promoter is needed to give rise to increased levels of transcription in response to auxin, but little expression in response to cytokinin, then the promoters of cDNAs that were up-regulated in the auxin experiments, but down-regulated the cytokinin experiments would be of

interest.

#### Time Course Experiments – Time Sensitive

Evaluation of time-course data as described above is also useful to identify time-specific  
5 promoters. Promoters or regulatory sequences, like the coding sequences, can be classified as early  
or late responding according to the microarray data. Promoters that facilitate expression of early or  
late genes are useful to direct expression of heterologous coding sequences to modulate the cellular  
response. In the drought data, promoters from “early” responding genes can be selected to activate  
expression of any desired coding sequence. Thus, a coding sequence for a salt-tolerance protein that  
10 is not typically expressed early in response to drought could be linked to an “early” responding  
promoter to increase salt tolerance within one hour after exposure to drought conditions.

#### Developmental Experiments – Time Sensitive

Another class of time-sensitive promoters and other regulatory sequence can be identified  
15 from the experiments examining different developmental stages. These regulatory sequences can  
drive transcription of heterologous sequence at particular times during development. For example,  
expression of stress-responsive genes during fruit development can protect any gain in fruit yield.

#### Common To Many Pathways – Cause General Effects

20 Promoters and other regulatory sequence associated with cDNAs that are differentially  
expressed in a number of similar responses can be used to cause general effects. These types of  
regulatory sequences can be used to inhibit or increase expression of a desired coding sequence in a  
number circumstances. For example, protein that is capable of acting as an insecticide can be  
placed under the control a general “stress” promoter to increase expression, not only when the plant  
25 is wounded, but under other stress attack.

## II. EXPERIMENTAL RESULTS ALSO REVEAL PATHWAYS OR NETWORKS OF GENES

### II.A. GENES WHOSE TRANSCRIPTION ARE WELL COORDINATED GENERALLY ACT TOGETHER TO PRODUCE PROTEINS THAT

PARTICIPATE IN THE SAME PATHWAY OR NETWORK

Patrick Brown, one of the pioneers of microarray chip technology, demonstrated that differential expression experiments can identify groups of genes that encode proteins that participate the same pathway or network. The work focused on phosphate accumulation and metabolism genes in yeast and was published in the paper Ogawa *et al.*, *Mol Biol Cell* (2000) Dec;11(12):4309-21. The authors identified by microarray analysis 22 genes whose transcription was regulated by phosphate concentration. Promoter analysis of these genes showed that 21 of them contained a sequence in their promoters that is recognized by a transcriptional activator that is regulated by phosphate. Further, phenotypic studies were completed by mutational analysis of many of these 22 genes in yeast. The mutants were shown to be either severely deficient in accumulation of inorganic polyphosphate (polyP) and P(i), or associated with normal catabolism of polyP in the yeast vacuole. This publication proves that genes with correlated transcriptional profiles do indeed participate in the same pathway or network.

II.A.1. CALCULATING THE CORRELATION COEFFICIENT BETWEEN  
PAIRS OF GENES BASED ON THE DIFFERENTIAL EXPRESSION  
DATA

The differential expression data obtained over many experiments reveal the global pattern of transcription of a gene. Transcription patterns, also referred to as profiles, of two different genes can be compared. From this comparison, a correlation coefficient can be calculated as a measure of the strength of the relationship between the two profiles.

Transcription profiles can be compared by plotting as a point, the differential expression of gene1 on the x-axis and gene 2 on the y-axis on one experiment. If all the pairs lie on a regression line the relationship and correlation between the two genes are strong. The correlation coefficient can be calculated using a number of methods. In the present case, the Spearman method was utilized.

The correlation coefficient can vary from -1 to 1. The coefficient indicates the strength of the relationship between two mRNA transcripts of any set of data that is examined. A zero coefficient indicates that no correlation exists between the transcription profiles of two genes in the samples examined.

Biologically, a high correlation coefficient indicates that a gene(s) triggers the activation or repression of the correlated genes, or have related functional roles. Thus, illumination of the activity of one gene can indicate the activities of the genes with highly correlated transcription profiles. This implication is true whether the activity is a biochemical activity, molecular interaction, cellular response, or physiological consequence.

#### II.A.2. THE COMPLETE LINKAGE ANALYSES OF DIFFERENTIAL IDENTITY GENES WITH SIMILAR PATTERN OF TRANSCRIPTION

The complete linkage analysis can build groups (or "clusters") of genes whose transcription patterns are highly correlated or co-regulated.

Because genes with related functions are frequently expressed in similar patterns, utilities or roles can be ascribed for genes (without observation of transformed plants) based on their temporal association with other genes of known function (a "guilt-by-association" analysis). Ogawa *et al.* has used correlated mRNA transcription profiles to identify the function of proteins of unknown function.

The complete linkage analysis utilizes the correlation coefficients that are calculated for each pair of genes tested in the microarray experiments. A cluster is first seeded with any arbitrary transcript tested on the chip. The seed transcript, for this illustration, is designated mRNA#0. Next, a minimum threshold is chosen for all acceptable correlation coefficients. In this case, the threshold used was 0.75. A list of potential cluster members is compiled by choosing mRNA transcripts that have a correlation coefficient with mRNA#0 that is greater than the threshold. No limit is placed on the number of mRNAs that can be added to a cluster so long as the correlation coefficient meets the threshold limit criterion.

For this example, assume that four mRNAs were added to the cluster, mRNA\_1 to mRNA\_4. Once the potential cluster members are identified, the cDNA IDs of each member is added to the potential list in order its correlation coefficient to mRNA#1, the largest correlation coefficient first. For this example, let's suppose four mRNAs 1-4 are potential members, they would be ordered as follows:

MRNA#	Correlation Coefficient with mRNA#0
-------	--

MRNA#1	0.9
MRNA#2	0.8
MRNA#3	0.78
MRNA#4	0.75

A potential member is accepted into the group, if its correlation coefficients with all other potential members are all greater than the threshold. Thus, for mRNA#1 to remain in the group the correlation coefficient between mRNA#1 and mRNA#2 must be greater than 0.75; and mRNA#1 and #3 > 0.75; and mRNA#1 and mRNA#4 > 0.75. Potential cluster members are removed only after reviewing the correlation coefficients in a specific order where mRNAs are reviewed in the order that they appear on the list.

Consequently, review of the correlation coefficients does not begin with any random pair, such as mRNA#3 and mRNA#4. The review begins between mRNA#1 and mRNA#2, which are the top two on the list.

If correlation coefficient between mRNA#1 and mRNA#2 is less than the threshold, then mRNA#2 is removed from the cluster. mRNA#2 is removed because its correlation coefficient with mRNA#0 is 0.8 which is less than 0.9, the correlation coefficient of mRNA#1 and mRNA#0.

This illustrates the rule that if the correlation coefficient is less than the threshold, then only one of the pair not accepted as a cluster member, specifically, the one with the lower coefficient to the seed mRNA#0.

This process of iterative reviewing of correlation coefficients between potential members continues until all pairs are reviewed. In this case, the coefficient between mRNA#1 and mRNA#3 would be reviewed because these are the two highest ones on the list besides mRNA#1 and #2. The next pair to be reviewed would be mRNA#1 and #4, etc.

Applicants have analyzed the data using several sets of parameters for the complete linkage analysis as shown in the table below:

Method	Correlation Coefficient Threshold	Max number of members in a cluster	Organism
CL_METHOD_TYPE=	0.9	MAX_SIZE=15	Arabidopsis

TRUE			
CL_METHOD_TYPE=TRUE	0.75	MAX_SIZE=30000	Arabidopsis
CL_METHOD_TYPE=TRUE	0.70	MAX_SIZE=30000	Arabidopsis
CL_METHOD_TYPE=TRUE	0.9	MAX_SIZE=15	Zea
CL_METHOD_TYPE=TRUE	0.75	MAX_SIZE=30000	Zea
CL_METHOD_TYPE=TRUE	0.70	MAX_SIZE=30000	Zea
CL_METHOD_TYPE=TRUE	0.9	MAX_SIZE=15	Arabidopsis
CL_METHOD_TYPE=TRUE	0.75	MAX_SIZE=30000	Arabidopsis
CL_METHOD_TYPE=TRUE	0.70	MAX_SIZE=30000	Arabidopsis
CL_METHOD_TYPE=TRUE	0.9	MAX_SIZE=15	Zea
CL_METHOD_TYPE=TRUE	0.75	MAX_SIZE=30000	Zea
CL_METHOD_TYPE=TRUE	0.70	MAX_SIZE=30000	Zea

The results of these cluster analyses are reported in the MA\_clust table.

### II.A.3. THE NEAREST NEIGHBOR ANALYSES OF DIFFERENTIAL GROUP GENES WITH CORRELATED BUT DISSIMILAR TRANSCRIPTION PROFILES

The nearest neighbor analysis differs from the complete linkage algorithm by not requiring all members to meet the correlation threshold with each other. Thus, a member of a nearest neighbor cluster need only be closely correlated to one other member of the cluster. It is not even required that all members be closely correlated to the seed mRNA transcript.

In a complete linkage cluster all the transcription profile of all members are correlated to a greater or lesser extent. In contrast, a cluster deduced by the nearest neighbor analysis may include members with differing transcription profiles. However, nearest neighbor brings to light

clusters of interacting genes. In the nearest neighbor analysis, the seed mRNA may not have a very high correlation coefficient with the last mRNA added to the cluster.

The nearest neighbor analysis, like the complete linkage analysis, is initiated by seeding each cluster with a mRNA\_0. The cluster size is determined by setting a threshold coefficient and setting a limit on the number of members that can be added to the cluster.

The cluster is expanded in an iterative fashion determining which mRNA has the highest correlation coefficient with mRNA\_0. The additional member is labeled mRNA\_1. Next, a list of potential candidates is generated by finding the mRNA that has the highest correlation to mRNA\_0 (besides mRNA\_1) and finding the mRNA that has the highest coefficient with mRNA\_1. Whichever of the candidates has the highest correlation coefficient is added to the cluster. Then, a list of three potential candidates is generated similarly.

Addition of members continues until either (1) all the correlation coefficients of potential members is lower than the threshold or (2) number of members in the cluster meets the size limitation.

Applicants have analyzed the data using several sets of parameters for the nearest neighbor analysis as shown in the table below:

Method	Correlation Coefficient Threshold	Max number of members in a cluster	Organism
NN_METHOD_TYPE=TRUE	0.5	MAX_HITS=15	Arabidopsis
FULL>NN_METHOD_TYPE=TRUE	0.8	NONE	Arabidopsis
FULL>NN_METHOD_TYPE=TRUE	0.6	NONE	Arabidopsis
NN_METHOD_TYPE=TRUE	0.5	MAX_HITS=15	Zea
FULL>NN_METHOD_TYPE=TRUE	0.8	NONE	Zea
FULL>NN_METHOD_TYPE=TRUE	0.6	NONE	Zea

NN_METHOD_TYPE=TRUE	0.5	MAX_HITS=15	Arabidopsis
FULL>NN_METHOD_TYPE=TRUE	0.8	NONE	Arabidopsis
FULL>NN_METHOD_TYPE=TRUE	0.6	NONE	Arabidopsis
NN_METHOD_TYPE=TRUE	0.5	MAX_HITS=15	Zea
FULL>NN_METHOD_TYPE=TRUE	0.8	NONE	Zea
FULL>NN_METHOD_TYPE=TRUE	0.6	NONE	Zea

The results of these cluster analyses are reported in the MA\_clust table.

### III. EXPERIMENTAL RESULTS REVEAL THE FUNCTIONS AND CHARACTERISTICS OF GENES, PATHWAYS AND NETWORKS

#### III.A. LINKING BIOCHEMICAL OR METABOLIC ACTIVITIES OF ONE PROTEIN IN A CLUSTER TO THE OTHER PROTEINS IN THE SAME MICROARRAY CLUSTER

As shown in the Ogawa *et al.*, *Mol Biol Cell* (2000), genes whose transcription profiles cluster together as being strongly correlated typically take part in the same pathway or network. Thus, the activity of one gene in the cluster can be associated to the other genes in the cluster with highly correlated transcription profiles. This association is true whether the activity is a biochemical activity, molecular interaction, cellular response or physiological consequence.

One example of this is cluster 420 of the report (shown below). In this cluster, a protein encoded by cDNA ID 1025791 did not match to any pFAM domain. However, through the microarray data, the gene that encodes that protein had a transcription profile that was correlated with other genes that encode ribosomal proteins. Thus, the activity of the ribosomal genes can be associated with the protein with no pFAM match. All the proteins in the same cluster would be associated with mRNA translation and protein synthesis.



420	1025791	803433	4585878	(AC005850) Unknown protein [Arabidopsis]		
420	4608965	671877	8567795	(AC013428) 40S ribosomal protein S17, pu	Ribosomal S17e	Ribosomal S17
420	5663116	818554	7486478	hypothetic al protein F6E13.17 - Arabidop	DapB	Dihydrodi picolinate reductase

5

### III.B. USING DIFFERENTIAL EXPRESSION DATA TO DETERMINE WHEN THE GENES AND PATHWAYS ARE ACTIVE

10 The differential expression data can be used to associate the cellular response that results  
when the clusters of genes are transcribed. For the complete linkage clusters, the genes in the  
cluster will produce similar transcription profiles. The experiments where the genes in the cluster  
are differentially expressed as compared to the control define the cellular responses that all the  
genes of the cluster are capable of modulating.

15 For example, for the cluster shown above, the mRNA levels for the genes were significantly  
different in the nitrogen response experiments. Thus, the data shows that this cluster of genes is  
associated with protein synthesis in response to nutrient uptake.

5            III.C. USING PHENOTYPE DATA TO DETERMINE WHEN GENES AND  
              PATHWAYS ARE ACTIVE

10            The phenotypic data can be used to demonstrate the physiological consequences of that  
15            result when a cluster of genes is active. Whether the clusters were generated by the complete  
              linkage or the nearest neighbor analyses, if a single gene in the cluster has been implicated in  
              phenotypic changes, then any one or combination of the other genes in the cluster can also modulate  
20            the same or similar phenotypic changes.

Utilities of Particular Interest

25            The following sections describes utilities/functions for the genes, gene components and  
              products of the invention. The sequences of the invention, as discussed above, can be recognized as  
              a particular type of gene (e.g. root gene, leaf gene, etc.) by means of particular terms utilized in the  
              Knock-in and Knock-out Tables and by the results of the differential expression experiments.  
              Combined analysis of those data also identify genes with utilities/functions of particular interest.  
              The Single Gene Functions and Utilities Table correlates that data and specific genes with those  
30            utilities/functions of particular interest.

Utilities of Particular Interest for Clustered Sequences

35            As discussed further herein, the genes, gene components and products of the invention have  
              been clustered together into groups. This enables one to understand the function/utility of one  
              member of the cluster based upon knowledge about one or more other members of the cluster. In  
              addition, this enables an understanding of some utilities/functions of a cluster that would be of  
              particular interest. The Cluster Functions and Utilities Table lists some of the clusters of the  
              invention and notes the functions/utilities that are of particular interest for each of the clusters. Of  
              course, these functions/utilities are of particular interest for each member of each particular cluster.

40            IV. EXPERIMENTAL RESULTS PROVIDE AN UNDERSTANDING OF GENES,

PATHWAYS AND NETWORKS IN MANY PLANT SPECIES

IV.A. INTRO: USING PROTEIN Sequence SIMILARITY TO ASSIGN FUNCTION  
FROM ARABIDOPSIS PROTEINS TO PROTEINS OF OTHER SPECIES

By analyzing the constant and variable properties of groups of similar sequences, it is possible to derive a structural and functional signature for a protein family, which distinguishes its members from all other proteins. This approach has allowed the Applicants to assign proteins into functional groups and identify orthologous proteins both within and between species. A pertinent analogy to be considered is the use of fingerprints by the police for identification purposes. A fingerprint is generally sufficient to identify a given individual. Similarly, a protein signature can be used to assign a newly sequenced protein to a specific family of proteins and thus to formulate hypotheses about its function.

Proteins can be grouped together because they share a single motif or many motifs. Typically, proteins that share a series of motifs share greater functional equivalence. Usually, signature sequences comprise more than one motif in a particular order from N-terminus to C-terminus.

A list of these groups can be found in the Protein Group Table. The sequences were grouped together using the iterative protein sequence local alignment software, PSI-BLAST. This software begins by aligning a number sequences where the probability that the alignment occurred by chance is set by a threshold e-value. In the Applicants' case, the threshold e-value was set at  $10^{-50}$ ,  $10^{-30}$ , and  $10^{-10}$ . The algorithm generates a consensus sequence from the sequences that were aligned together. The consensus sequence was then used to find sequences that matched to it with a probability that was less than the set threshold. The algorithm performs the iterative tasks of aligning and generating a consensus sequence any number of times. Generally, Applicants performed one iteration for the  $10^{-10}$  e-value threshold, two iterations for the  $10^{-30}$  threshold, and three iterations for the  $10^{-50}$  threshold.

Each group can contain sequences from one of more organisms. The groups included both Ceres polypeptides and public polypeptide sequences. The Ceres polypeptides are identified by their Ceres Sequence ID NO as listed in the Reference Table.

Each group contains sequences that were included at the  $10^{-50}$ ,  $10^{-30}$ , and  $10^{-10}$  e-value cutoffs. For each group, the peptide ID and at which cutoff the peptide was included into the group.

The same peptide ID may be included in the group three times as peptide ID 50, peptide ID 30 and peptide ID 10. The data indicates that peptide ID was included in the group when the threshold was either  $10^{-50}$ ,  $10^{-30}$ , or  $10^{-10}$ . All the peptide IDs that are followed by "50" were included in the protein group when the e-value cutoff was  $10^{-50}$ . All the peptide IDs that are followed by either "30" or "50" were included in the protein group when the threshold e-value was  $10^{-30}$ . All the peptide IDs that are followed by "10", "30" or "50" were included in the protein group when  $10^{-10}$  was used as the e-value cutoff.

#### IV.A.1. CONSERVED SEQUENCES BETWEEN PROTEINS OF DIFFERENT SPECIES GIVE RISE TO A SIGNATURE SEQUENCE

The signature sequence for each group of proteins, also referred to as the consensus sequence. The signature sequence comprises the amino acids that are conserved throughout all the proteins in a particular protein group. The data are shown in the Protein Group table.

Not all the polypeptides in a group are the same length. Thus, some members of the group may not contain the entire signature sequence. However, throughout the length of any member protein, its sequence will match the signature sequence.

The consensus sequence contains both lower-case and upper-case letters. The upper-case letters represent the standard one-letter amino acid abbreviations. The lower case letters represent classes of amino acids:

- "t" refers to tiny amino acids, which are specifically alanine, glycine, serine and threonine.
- "p" refers to polar amino acids, which are specifically, asparagine and glutamine
- "n" refers to negatively charged amino acids, which are specifically, aspartic acid and glutamic acid
- "+" refers to positively charged residues, which are specifically, lysine, arginine, and histidine
- "r" refers to aromatic residues, which are specifically, phenylalanine, tyrosine, and tryptophan,
- "a" refers to aliphatic residues, which are specifically, isoleucine, valine, leucine, and methionine

In addition to each consensus sequence, Applicants have generated a scoring matrix to provide further description of the consensus sequence. The matrix reports the identity and number of occurrences of all the amino acids that were found in the group members for every residue position of the signature sequence. The matrix also indicates for each residue position, how many different organisms were found to have a polypeptide in the group that included a residue at the relevant position. These results are reported in the Protein Group Matrix table.

Functional equivalents share similar (1) structural characteristics; (2) biochemical activities and molecular interactions; (3) cellular responses or activities; or (4) phenotypic effects.

#### IV.A.2.LINKING SIGNATURE SEQUENCES TO CONSERVATION OF STRUCTURAL CHARACTERISTICS

Proteins with related functions show similar three-dimensional structures but may not show extensive amino acid sequence similarity. Typically, proteins need only share a single motif or low similarity in multiple domains to exhibit similar structural features, such as alpha helix, beta sheet, charge residues, stretches of hydrophobicity, etc. Conserved structural features have been implicated in ligand binding by receptor proteins, binding to a class of substrates, polynucleotide binding, or protein-protein interactions.

Based on the signature sequences and the Matrix Tables described herein, a number of motifs can be discerned. Motifs are identified as regions in the signature sequence which are constant in a majority of the members of the group. Example motifs can be found among Applicant's data which are shared in the range of 75% to 95% of group members

Typically, a region of the consensus sequence is constant if, at each position of the region, the preferred amino acid is chosen from a single class of amino acids; even more typically, the preferred amino acid is a single amino acid. The region can contain a number of positions where an amino acid can be chosen. However, these variable positions are usually less than 15% of the total number of residues in the region; more usually, less than 10%; even more usually, less than 5%.

Generally, a domain is considered to be well defined if the consensus sequence is constructed from sequences from at least 2 organisms; more preferably, at least 3 organisms; even more preferably four organisms or greater.

Primary domains are best identified from the data presented for the  $10^{-10}$  probability criteria.

Using this parameter, the largest number of proteins is associated into a group. Consequently, the signature sequence exhibits the greatest amount of variability. The conserved regions, the domains or motifs of the signature contrast against the variable regions. These variable regions become obvious when sequences from more proteins are compared.

Signature sequences revealed in the  $10^{-30}$  and  $10^{-50}$  e-value classes show more conservation in the domains, and can even display a degree of conservation in what is considered the variable regions in the  $10^{-10}$  analyses. These more extensively-conserved domains can reflect higher similarity in function – completely orthologous functions. Proteins that share a number of conserved domains, in the same relative order from N terminus to C terminus, are even more likely to be completely orthologous. Nevertheless, because of the natural divergence that occurs in non-conserved regions during evolution and species differentiation, orthologs can be proteins with only the domains conserved and therefore be present in the  $10^{-30}$  and  $10^{-10}$  p value classes of the Ortholog Table.

#### IV.A.3.LINKING SIGNATURE SEQUENCES TO CONSERVATION OF BIOCHEMICAL ACTIVITIES AND MOLECULAR INTERACTIONS

Proteins that possess the same defined domains or motifs are likely to carry out the same biochemical activity or interact with a similar class of target molecule, e.g., DNA, RNA, proteins, etc. Thus, the pFAM domains listed in the Reference Tables are routinely used as predictors of these properties. Substrates and products for the specific reactions can vary from protein to protein. Where the substrates, ligands, or other molecules bound are identical the affinities may differ between the proteins. Typically, the affinities exhibited by different functional equivalents varies no more than 50%; more typically, no more than 25%; even more typically, no more than 10%; or even less.

Proteins with very similar biochemical activities or molecular interactions will share similar structural properties, such as substrate grooves, as well as sequence similarity in more than one motif. Usually, the proteins will share at least two motifs of the signature sequence; more usually, three motifs; even more usually four motifs or greater. Typically, the proteins exhibit 70% sequence identity in the shared motifs; more typically, 80% sequence identity; even more typically, 90% sequence identity or greater. These proteins also often share sequence similarity in the variable regions between the constant motif regions. Further, the shared motifs will be in the same order

from amino- to carboxyl-termini. The length of the variable regions between the motifs in these proteins, generally, is similar. Specifically, the number of residues between the shared motifs in these proteins varies by less than 25%; more usually, does not vary by less than 20%; even more usually, less than 15%; even more usually less than 10% or even less.

#### IV.A.4.LINKING SIGNATURE SEQUENCES TO CONSERVATION OF CELLULAR RESPONSES OR ACTIVITIES

Proteins that exhibit similar cellular response or activities will possess the structural and conserved domain/motifs as described in the Biochemical Activities and Molecular Interactions above.

Proteins can play a larger role in cellular response than just their biochemical activities or molecular interactions suggest. A protein can initiate gene transcription, which is specific to the drought response of a cell. Other cellular responses and activities include: stress responses, hormonal responses, growth and differential of a cell, cell to cell interactions, etc.

The cellular role or activities of protein can be deduced by transcriptional analyses or phenotypic analyses as well as by determining the biochemical activities and molecular interactions of the protein. For example, transcriptional analyses can indicate that transcription of gene A is greatly increased during flower development. Such data would implicate protein A encoded by gene A, in the process of flower development. Proteins that shared sequence similarity in more than one motif would also act as functional equivalents for protein A during flower development.

#### IV.A.5.LINKING SIGNATURE SEQUENCES TO CONSERVATION OF PHENOTYPIC EFFECTS

Typically, proteins that are grouped together under the most stringent parameters, e-value  $\leq 10^{-50}$ , are likely orthologs and therefore, when present in the same or equivalent cells can cause similar phenotypic consequences. These proteins have very high sequence similarity. Typically, if one of the members of a group is an *Arabidopsis* protein, then the corn ortholog can rescue an *Arabidopsis* mutant plant that does not produce the *Arabidopsis* protein. The mutant plant would be rescued as the parental "wild-type" phenotype by expression of a coding sequence of the corn protein of the same orthologous group when present in the appropriate cell types of the plant.

Preferably, these functional equivalents have sequence motif identity throughout much of

the length of the protein. However, proteins that share very high similarity between a number, usually more than two; even more usually, more than three motifs can act as functional equivalents to produce similar phenotypic effects.

5 A gene can have coding sequence similarity, i.e., is a homologous. The coding sequence can be sufficient to act as a functional equivalent, although the gene as a whole is not an ortholog. For example, two similar *dwf4* coding sequences were found in the *Arabidopsis* genome. However, this pair of coding sequences had different promoters and hence different roles in *Plantae*. But when one of the pair was placed under the control of its mates' promoter, the phenotypic effects were similar to the effects produced by its mate coding sequence. Therefore, the coding sequence, but not the genes are orthologous.

10



## APPLICATIONS OF THE INVENTIONS

As described herein, the results of Applicant's experiments provide an understanding of the function and phenotypic implications of the genes, gene components and products of the present invention. Bioinformatic analysis provides such information. The sections of the present application containing the bioinformatic analysis, together with the Sequence and Reference Tables, teach those skilled in the art how to use the genes, gene components and products of the present invention to provide plants with novel characteristics. Similarly, differential expression analysis provides additional such information and the sections of the present application on that analysis; together with the MA\_diff and/or AFLP\_diff Tables and MA\_Cluster Tables, describe the functions of the genes, gene components and products of the present invention which are understood from the results of the differential expression experiments. The same is true with respect to the phenotype data, wherein the results of the Knock-in and Knock-out experiments and the sections of the present application on those experiments provide the skilled artisan with further description of the functions of the genes, gene components and products of the present invention.

As a result, one reading each of these sections of the present application as an independent report will understand the function of the genes, gene components and products of the present invention. But those sections and descriptions can also be read in combination, in an integrated manner, to gain further insight into the functions and uses for the genes, gene components and products of the present invention. Such an integrated analysis does not require extending beyond the teachings of the present application, but rather combining and integrating the teachings depending upon the particular purpose of the reader.

Some sections of the present application describe the function of genes, gene components and products of the present invention with reference to the type of plant tissue (e.g. root genes, leaf genes, etc.), while other sections describe the function of the genes, gene components and products with respect to responses under certain conditions (e.g. auxin-responsive genes, heat-responsive genes, etc.). Thus, if one desires to utilize a gene understood from the application to be a particular tissue-type of gene, then the condition-specific responsiveness of that gene can be understood from the differential expression tables, and very specific characteristics of actions of that gene in a transformed plant will be understood by recognizing the overlap or intersection of the gene

functions as understood from the two different types of information. Thus, for example, if one desires to transform a plant with a root gene for enhancing root growth and performance, one can know the useful root genes from the results reported in the knock-in and knock-out tables. A review of the differential expression data may then show that a specific root gene is also over-expressed in response to heat and osmotic stress. The function of that gene is then described in (1) the section of the present application that discusses root genes, (2) the section of the present application that discusses heat-responsive genes, and (3) the section of the application that discusses osmotic stress-responsive genes. The function(s) which are commonly described in those three sections will then be particularly characteristic of a plant transformed with that gene. This type of integrated analysis of data can be viewed from the following schematic that summarizes, for one particular gene, the function of that gene as understood from the phenotype and differential expression experiments.

Gene function known from phenotype experiments	Gene function known from first differential expression experiment	Gene function known from second differential expression experiment
Function A	Function A	Function A
Function B		
	Function C	Function C
	Function D	
		Function E
Function F	Function F	Function F
Function G	Function G	
		Function H
Function I		Function I
	Function J	

In the above example, one skilled in the art will understand that a plant transformed with this particular gene will particularly exhibit functions A and F because those are the functions which are

understood in common from the three different experiments.

Similar analyses can be conducted on various genes of the present invention, by which one skilled in the art can effectively modulate plant functions depending upon the particular use or conditions envisioned for the plant.

I. ORGAN-AFFECTING GENES, GENE COMPONENTS, PRODUCTS (INCLUDING DIFFERENTIATION AND FUNCTION)

I.A. ROOT GENES, GENE COMPONENTS AND PRODUCTS

I.A.1. ROOT GENES, GENE COMPONENTS AND PRODUCTS

The economic values of roots arise not only from harvested adventitious roots or tubers, but also from the ability of roots to funnel nutrients to support growth of all plants and increase their vegetative material, seeds, fruits, etc. Roots have four main functions. First, they anchor the plant in the soil. Second, they facilitate and regulate the molecular signals and molecular traffic between the plant, soil, and soil fauna. Third, the root provides a plant with nutrients gained from the soil or growth medium. Fourth, they condition local soil chemical and physical properties.

a. Identification of Root Genes

Root genes identified herein are defined as genes, gene components and products capable of modulating one or more processes in or functions of the root as described below. They are active or potentially active to a greater extent in roots than in most other organs of the plant.

These genes and gene products can regulate many plant traits from yield to stress tolerance. That single genes usually affect the development and function of roots and whole plants is a consequence of biological cellular complexity and the role roots play in supporting the growth of whole plants. Examples of such root genes and gene products are shown in the Reference and Sequence Reference and Sequence Tables and sequences encoding polypeptides of the Protein Group and Protein Group Matrix tables or fragments thereof, the Knock-In and Knock-Out Tables, and the MA-diff Tables. The function of many of the protein products gained from comparisons with proteins of known functions, are also given in the REF Tables.

Root Genes Identified By Phenotypic Observations

Root genes are active or potentially active to a greater extent in roots than in some other organs/tissue of the plant. Some of the root genes herein were discovered and characterized from

a much larger set of genes in experiments designed to find genes that cause phenotypic changes in root morphology. Such morphological changes include primary and lateral root number, size and length, as well as phenotypic changes of other parts of that plant associated with changes in root morphology.

In these experiments, root genes were identified by either (1) ectopic expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then cultivated under standardized conditions and any phenotypic differences recorded between the modified plants as compared with the parent plant. The gene(s) causing the changes were deduced from the cDNA inserted or disrupted gene. Phenotypic differences were observed in:

Primary Roots And Root System

- Size, Including Length And Girth
- Number
- Branching
- Root Waving/Curling Characteristics
- Gravitropism Changes
- Agravitropic

Lateral Roots

- Size, Including Length And Girth
- Number
- Branching

Results from screening for these phenotypic changes are reported in the Knock-in and Knock-out Tables. Therefore, any sequence reported in those Tables with one of the above-noted observations is considered a “root gene”. A “root gene” is also a sequence which, in the Ortholog Tables or in the MA-clust Tables, is grouped/clustered together with at least one sequence that is identified as such by means of the Knock-in and Knock-out Tables.

Root Genes Identified By Differential Expression

Root genes were also identified by measuring the relative levels of mRNA products in the root versus the aerial portion of a plant. Specifically, mRNA was isolated from roots and root tips

of Arabidopsis plants and compared to mRNA isolated from the aerial portion of the plants utilizing microarray procedures. The MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108594, 108433, 108599, 108434, 108439 ). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Roots genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication or that are shown as having differential expression in the AFLP\_diff table.

#### Roots Genes Identified By Cluster Analyses Of Differential Expression

##### Roots Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Roots genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108594, 108433, 108599, 108434, 108439 of the MA\_diff and/or AFLP\_diff table(s).

##### Roots Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Roots genes. A group in the MA\_clust is considered a Roots pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Roots Genes Identified By Amino Acid Sequence Similarity

Roots genes from other plant species typically encode polypeptides that share amino acid

similarity to the sequences encoded by corn and Arabidopsis Roots genes. Groups of Roots genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Roots pathway or network is a group of proteins that also exhibits Roots functions/utilities.

Examples of phenotypes, biochemical activities, and transcription profiles that can be modulated by these genes and gene products are described above and below.

b. Identification of Root Genes

b. Use of Root Genes to Modulate Phenotypes

The root genes of the instant invention are capable of modulating one or more processes of root structure and/or function including (I) development; (II) interaction with the soil and soil contents; and (III) transport in the plant.

Root genes and gene products can be used to alter or modulate one or more of the following phenotypes.

I. DEVELOPMENT

Roots arise from meristem cells that are protected by a root cap during root elongation, but as the root grows out, the cap cells abscise and the remaining cells differentiate to the tip. Depending on the plant species, some surface cells of roots can develop into root hairs. Some roots persist for the life of the plant; others gradually shorten as the ends slowly die back; some may cease to function due to external influences. The root genes and gene products of this invention are useful to modulate any one or all of these growth and development processes generally, as in root density and root growth; including rate, timing, direction, size, for example.

A. Different Types Of Roots

Root genes and gene products are useful to modulate either the growth and development or other processes in one or more of the following types of roots:

1. Primary
2. Lateral
3. Root Hairs (See The Section Below For More Detail)
4. Adventitious

B. Cell Properties

Root genes and gene products are useful to modulate cellular changes in:

1. Cell Size
2. Cell Division, Rate Direction And/Or Number
3. Cell Elongation
4. Cell Differentiation
5. Lignified Cell Walls
6. Epidermal Cells, Such As Trichoblasts
7. Root Apical Meristem Cells (Growth And Initiation)

C. Root Architecture

The following parts of roots can be modulated by these genes root and gene products to affect root architecture:

1. Epidermis
2. Cortex
  - (a) Epidermis
  - (b) Hypodermis
  - (c) Endodermis
  - (d) Casparian Strips
  - (e) Suberized Secondary Walls
  - (f) Parenchyma
  - (g) Aerenchyma
3. Stele
  - (a) Vacuature



- (i) Xylem
- (ii) Phloem
- (b) Pericycle

- 4. Vasculture
- 5. Xylem
- 6. Phloem
- 7. Root Cap
- 8. Root Apical Meristem
- 9. Elongating Region
- 10. Symmetry

D. Root Responses

The polynucleotides and polypeptides of this invention can be used to control the responses to internal plant programs as well as to environmental stimuli in:

- 1. Seminal System
- 2. Nodal System
- 3. Hormone Systems
  - (a) Auxin (For More Details See The Section On Auxin-Responsive Genes)
  - (b) Cytokinin (Inhibitory For Root Development, See Section On Cytokinin-Responsive Genes)
- 4. Root Cap Abcission
- 5. Root Senescence
- 6. Gravitropism
- 7. Coordination Of Root Growth And Development With That Of Other Organs, Examples Include:
  - (a) Leaves
  - (b) Flowers
  - (c) Seeds
  - (d) Fruits
  - (e) Stems
- 8. Changes In Soil Environment (For More Detail See Below)

- (a) Water
- (b) Minerals
- (c) Ph
- (d) Microfauna And Flora

5

## II. INTERACTION WITH SOIL AND SOIL CONTENTS

Roots are sites of intense chemical and biological activities and as a result can strongly modify the soil they contact. Roots coat themselves with surfactants and mucilage to facilitate these activities. Specifically, roots are responsible for nutrient uptake by mobilizing and assimilating water, organic and inorganic compounds, ions and attracting and interacting with beneficial microfauna and flora. Roots also help to mitigate the effects of toxic chemicals, pathogens and stress. Examples of root properties and activities that the genes and gene products of this invention are useful to modulate are as follows:

10  
15  
20  
25  
30

- A. Root Surfactants And Mucilage
  - 1. Mucilage
    - (a) Composition
    - (b) Secretion Rate And Time
  - 2. Surfactant
- B. Nutrient Uptake Of
  - 1. Water
    - (a) Which Can Be Measured By The Supply To Shoot On The Basis Of Volume/Dry Weight Or Surface Area
  - 2. Nitrate And Other Sources Of Nitrogen (For Detail See The Section On Nitrogen Responsive Genes)
  - 3. Phosphate
  - 4. Potassium
  - 5. Micronutrients (E.G. Iron, Copper, Etc.)
- C. Microbes And Nematodes Associations:
  - 1. Bacteria Including Nitrogen-Fixing Bacteria
  - 2. Mycorrhizae
  - 3. Nodule-Forming And Other Nematodes

- D. Oxygen
  - 1. Transpiration
- E. Detoxification Of
  - 1. Iron
  - 2. Aluminum
  - 3. Cadmium
  - 4. Mercury
  - 5. Salt
  - 6. Other Heavy Metals And Toxins
- F. Pathogen Interactions/Control
  - 1. Chemical Repellents
    - (a) Glucosinolates (GSL), Which Release Pathogen-Controlling Isothiocyanates
- G. Changes In Soil Properties, Such As:
  - 1. Ph
  - 2. Mineral Depletion
  - 3. Rhizosphere

### III. TRANSPORT OF MATERIALS IN PLANTS

Uptake of nutrients by roots produces a “source-sink” effect in a plant. The greater the source of nutrients, the larger “sinks,” such as stems, leaves, flowers, seeds, fruits, etc. can grow. Thus, root genes and gene products are useful to modulate the vigor and yield of the plant overall as well as distinct cells, organs, or tissues. Examples are as follows:

- A. Vigor
  - 1. Plant Nutrition
  - 2. Growth Rate
    - Whole Plant, Including Height, Flowering Time, Etc.
    - Seedling
    - Coleoptile Elongation
    - Young Leaves

- Stems
  - Flowers
  - Seeds
  - Fruit
3. Yield
- (A) Biomass
    - Fresh And Dry Weight During Any Time In Plant Life, Including Maturation And Senescence
  - (b) Root/Tuber Yield
    - Number, Size, Weight, Harvest Index
    - Content And Composition, E.G. Amino Acid, Jasmonate, Oil, Protein And Starch
  - (c) Number Of Flowers
  - (d) Seed Yield
    - Number, Size, Weight, Harvest Index
    - Content And Composition, E.G. Amino Acid, Jasmonate, Oil, Protein And Starch
  - (e) Fruit Yield
    - Number, Size, Weight, Harvest Index, Post Harvest Quality
    - Content And Composition, E.G. Amino Acid, Jasmonate, Oil, Protein And Starch

Additional Uses Of Plants With Modified Roots

Plants with roots modified in one or more of the properties described above in I, II or III are used to provide:

- A. Higher vigor and yield of plants and harvested products due to pathogen resistance from conditioning the soil with plant-derived chemicals and/or more tolerance to stresses such as drought, flooding and anoxia.
- B. Better Animal (Including Human) Nutrition
- C. Improved Dietary Mineral Nutrition

- D. Better Plant Survival
  - (a) Decreased Lodging
  - (b) More Efficient Transport
  - (c) More Efficient Physiology
  - (d) More Efficient Metabolism
- E. Better Resistance To Plant Density Effects
- F. Increased Yield Of Valuable Molecules
- G. More Efficient Root Nodulation
- H. Better Access To Rhizobia Spray Application, For Anaerobic Soils
- I. Easier Crop Harvesting And Ground Tillage
- J. Decreased Soil Erosion

To regulate any of the phenotype(s) above, activities of one or more of the root genes or gene products is modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Dolan et al. (1993, Development 119: 71-84), Dolan et al. (1997, Development 124: 1789-98), Crawford and Glass (1998, Trends Plant Science 3: 389-95), Wang et al. (1998, PNAS USA 95: 15134-39), Gaxiola et al. (1998, PNAS USA 95: 4046-50), Apse et al. (1999, Science 285: 1256-58), Fisher and Long (1992, Nature 357: 655-60), Schneider et al. (1998, Genes Devel 12: 2013-21) and Hirsch (1999, Curr Opin Plant Biol. 2: 320-326).

- c. Use of Root Genes to Modulate Phenotypes
- c. Use of Root Genes to Modulate Biochemical Activities

The activities of one or more of the root genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Association Of Root Morphology With Nitrogen Fixing Bacteria	<ul style="list-style-type: none"> <li>Cell-Cell Recognition</li> <li>Cell Wall Degradation</li> </ul>	Gage et al. (1996) J Bacteriol 178: 7159-66
Primary Root, Lateral Root, And Root Hair <ul style="list-style-type: none"> <li>Initiation</li> <li>Spacing</li> <li>Elongation</li> <li>Branching</li> </ul>	<ul style="list-style-type: none"> <li>Cell Division/Elongation</li> <li>Cell Differentiation</li> <li>Cell Expansion</li> <li>Auxin Mediated Response Pathways</li> </ul>	Schneider et al. (1998) Genes Devel 12: 2013-21 Casimiro et al. (2001). Plant Cell 13:843-852. Rogg et al. (2001). Plant Cell 13:465-480. Gaedeke et al. (2001). EMBO J. 20:1875-1887. Neuteboom et al. (1999). Plant Mol. Biol. 39:273-287. Schindelman et al. (2001). Genes and Dev. 15:1115- 1127. Rashotte et al. (2001) Plant Cell 13:1683-1697. Zhang et al. (2000). J Exp Bot 51:51-59. Zhang et al. (1998) Science 279: 407-409.
Metabolism	<ul style="list-style-type: none"> <li>Organic Molecule Export</li> </ul>	Moody et al. (1988) Phytochemistry 27: 2857-61.
	<ul style="list-style-type: none"> <li>Ion Export</li> </ul>	Uozumi et al. (2000) Plant Physiol 122: 1249-59 Frachisse et al. (2000) Plant J 21: 361-71

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	<ul style="list-style-type: none"> <li>Nutrient Uptake</li> </ul>	<p>Frachisse et al. (2000) Plant J 21: 361-71</p> <p>Uozumio et al. (2000) Plant Physiol 122: 1249-59</p> <p>Williamson et al. (2001). Plant Physiol. 126:875-882.</p> <p>Zhang et al. (2000). J Exp Bot 51:51-59.</p> <p>Zhang et al. (1998). Science 279:4 07-409.</p> <p>Coruzzi et al. (2001). Plant Physiol. 125: 61-64.</p>
Root Gravitropism And Waving	<ul style="list-style-type: none"> <li>Reactive Oxygen Species (ROS) Such As Superoxide Anions And H<sub>2</sub>O<sub>2</sub> Production</li> <li>Auxin Transport Pathways</li> <li>Flavonoid Inhibition Of Auxin Transport Function</li> <li>Changes In Root Cap Ph</li> <li>Starch Synthesis And Storage</li> <li>Cell Differentiation</li> <li>Cell Elongation</li> </ul>	<p>Joo et al. (2001) Plant Physiol. 126:1055-60.</p> <p>Vitha et al. (2000). Plant Physiol. 122: 453-461.</p> <p>Tasaka et al. (2001) Int Rev Cytol 206:135-54.</p> <p>Brown et al. (2001) Plant Physiol 126:524-35.</p> <p>Fasano et al. (2001) Plant Cell 13:907-22.</p> <p>MacCleery et al. (1999). Plant Physiol 120:183-92</p> <p>Blancaflor et al. (1998). Plant Physiol 116:213-22</p> <p>Schneider et al. (1998) Genes Devel 12: 2013-21</p>

Other biological activities that can be modulated by the root genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

A.1.a. USE OF ROOT GENES TO MODULATE TRANSCRIPTION  
LEVELS OF PLANT GENES

Many genes are “up regulated “ or “down regulated” because they belong to networks or cascades of genes. Thus some root genes are capable of regulating many other gene activities via these networks and hence complex phenotypes. Examples of transcription profiles of root genes are described in the Table below with associated biological activities. “Up-regulated” profiles are those where the concentrations of the mRNA in total mRNA are higher in roots as compared to aerial parts of a plant; and vice-versa for “down-regulated” profiles.



TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts	<p>Genes Expressed In Root Development</p> <p>Responders To Micro-Organismal Symbionts And Parasites</p> <p>Genes involved in polar auxin transport</p> <p>Genes involved in starch deposition in the roots</p> <p>Genes involved in production of reactive oxygen species</p> <p>Genes involved in</p>	<ul style="list-style-type: none"> <li>• Primary Root, Lateral Root, and/or Root Hair Growth and Differentiation</li> <li>• Microorganism Perception</li> <li>• Entrapment Of Microorganismal Symbionts</li> <li>• Nutrient Uptake</li> <li>• Synthesis Of Metabolites And/Or Proteins</li> <li>• Modulation Of Transduction Pathways</li> <li>• Specific Gene Transcription Initiation</li> <li>• Nutrient Uptake Enhancement</li> <li>• Gravitropic growth of roots</li> <li>• Associations with rhizobia are stimulated</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic Enzymes</li> <li>• Change In Cell Membrane Structure And Potential</li> <li>• Kinases, Phosphatases, G-Proteins</li> <li>• Transcription Activators</li> <li>• Change In Chromatin Structure And/Or Localized DNA Topology</li> <li>• Cell Wall Proteins</li> <li>• <math>Ca^{++}</math> Fluctuation</li> <li>• Reactive Oxygen Species (ROS) production</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	flavonoid synthesis		
Down-Regulated Transcripts	<p>Genes Repressed In Root Development</p> <p>Responders To Micro-Organismal Symbionts And Parasites</p> <p>Genes With Discontinued Expression Or UnsTable mRNA In Presence Of Root And/Or Micro-Organismal Symbionts</p>	<ul style="list-style-type: none"> <li>• Negative Regulation Of Primary Root, Lateral Root, and/or Root Hair Production Released</li> <li>• Changes In Pathways And Processes Operating In Cells</li> <li>• Changes In Metabolism</li> <li>• Inhibition of root gravitropism</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Kinases, Phosphatases, G-Proteins</li> <li>• Change In Chromatin Structure And/Or DNA Topology</li> <li>• Stability Of Factors For Protein Synthesis And Degradation</li> <li>• Metabolic Enzymes</li> </ul>

Changes in the function or development of roots are the result of modulation of the activities of one or more of these many root genes and gene products. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield, especially when plants are growing in the presence of soil borne biotic or abiotic stresses or when they are growing in barren conditions or in soils depleted of certain minerals.

Root genes, gene components and gene products can act alone or in combination as described in the introduction. Of particular interest are combinations of these genes and gene products with those that modulate stress tolerance and/or metabolism. Stress tolerance and metabolism genes and gene products are described in more detail in the sections below.

#### Use of Promoters of Root Genes

Promoters of root genes, as described in the Reference tables, for example, can be used to modulate transcription that is induced by root development or any of the root biological processes or activities above. For example, when a selected polynucleotide sequence is operably linked to a promoter of a root gene, then the selected sequence is transcribed in the same or similar temporal, development or environmentally-specific patterns as the root gene from which the promoter was taken. The root promoters can also be used to activate antisense copies of any coding sequence to achieve down regulation of its protein product in roots. They can also be used to activate sense copies of mRNAs by RNA interference or sense suppression in roots.

I.A.2. ROOT HAIR GENES, GENE COMPONENTS AND PRODUCTS

Root hairs are specialized outgrowths of single epidermal cells termed trichoblasts. In many and perhaps all species of plants, the trichoblasts are regularly arranged around the perimeter of the root. In Arabidopsis, for example, trichoblasts tend to alternate with non-hair cells or atrichoblasts. This spatial patterning of the root epidermis is under genetic control, and a variety of mutants have been isolated in which this spacing is altered or in which root hairs are completely absent.

a) IDENTIFICATION OF ROOT HAIR GENES

Root hair genes identified herein are defined as genes, gene components and products capable of modulating one or more processes in or the function of root hairs as described below. Root hairs are capable of controlling or influencing many plant traits, also as shown below. Examples of such root hair development genes and gene products are shown in the Reference and Sequence Tables. The protein products of many of these genes are also identified in these Tables.

Root Hair Genes Identified by Differential Expression

These genes were discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA products are associated specifically with root hairs. These experiments made use of the arabidopsis mutant "root hairless" (rhl), which does not develop root hairs. By comparing gene expression profiles of rhl roots with those of wild type roots grown in identical conditions, genes specifically expressed in root hairs were revealed. The MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108594, 108433). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Root Hairs genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication

Root Hairs Genes Identified By Cluster Analyses Of Differential Expression

Root Hairs Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Root Hairs genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108594, 108433 the MA\_diff and/or AFLP\_diff table(s).

Root Hairs Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Root Hairs genes. A group in the MA\_clust is considered a Root Hairs pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Root Hairs Genes Identified By Amino Acid Sequence Similarity

Root Hairs genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Root Hairs genes. Groups of Root Hairs genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Root Hairs pathway or network is a group of proteins that also exhibits Root Hairs functions/utilities.

Examples of phenotypes, biochemical activities, and transcript profiles that can be modulated by these genes and gene products are described above and below.

b) USE OF ROOT HAIR DEVELOPMENT GENES TO MODULATE PHENOTYPES

The root hair development genes of the instant invention are useful to modulate one or more processes of root hair structure and/or function including (I) development; (II) interaction

with the soil and soil contents; (III) uptake and transport in the plant; and (IV) interaction with microorganisms.

5

## I. DEVELOPMENT

10

The surface cells of roots can develop into single epidermal cells termed trichoblasts or root hairs. Some of the root hairs will persist for the life of the plant; others will gradually die back; some may cease to function due to external influences. The genes and gene products of this invention are useful to modulate any one or all of these growth and development processes generally, as in root hair density or root hair growth; including rate, timing, direction, and size, for example. A more detailed list of processes that are regulated by these genes and gene products is as follows:

15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995

### A. Cell Properties

Root hair development genes and gene products modulate cellular changes in:

1. Cell Size
2. Cell Division, Rate And Direction And Number
3. Cell Elongation
4. Cell Differentiation
5. Lignified Cell Walls
6. Epidermal Cells, Such As Trichoblasts
7. Root Apical Meristem Cells (Growth And Initiation)

25

### B. Root Hair Architecture

The following parts of a root hair can be modulated to affect root hair architecture:

- A. Leaf Cells Under The Trichome
- B. Cells Forming The Base Of The Trichome
- C. Trichome Cells

30

F. Root Hair Responses

The genes and gene products of this invention are useful to modulate any one or all of the following growth and development processes in response to internal plant programs or environmental stimuli in:

1. Seminal System
2. Nodal System
3. Hormone Responses
4. Auxin (For More Details See The Section On Auxin-Responsive Genes)
5. Root Cap Abscission
6. Root Senescence
7. Gravitropism
8. Coordination Of Root Growth And Development With That Of Other Organs, Examples Include:
  - (a) Leaves
  - (b) Flowers
  - (c) Seeds
  - (d) Fruits
  - (e) Stems
9. Changes In Soil Environment (For More Detail See Below)
  - (a) Water
  - (b) Minerals
  - (c) Ph
  - (d) Microfauna And Flora

II. INTERACTION WITH SOIL AND SOIL CONTENTS

Root hairs are sites of intense chemical and biological activity and as a result can strongly modify the soil they contact. Roots hairs can be coated with surfactants and mucilage to facilitate these activities. Specifically, roots hairs are responsible for nutrient uptake by mobilizing and assimilating water, reluctant ions, organic and inorganic compounds and

chemicals. In addition, they attract and interact with beneficial microfauna and flora. Root hairs also help to mitigate the effects of toxic ions, pathogens and stress. Examples of root hair properties and activities that the genes and gene products of the invention are useful to modulate are as follows:

- A. Root Hair Surfactant And Mucilage
  1. Mucilage
    - (a) Composition
    - (b) Secretion Rate And Time
  2. Surfactant
- B. Nutrient Uptake
  1. Water
  2. Nitrate And Other Sources Of Nitrogen(For Detail See The Section On Nitrogen Responsive Genes)
  3. Phosphate
  4. Potassium
  5. Micronutrients (E.G. Iron, Copper, Etc.)
- C. Microbe And Nematode Associations
  1. Bacteria Including Nitrogen-Fixing Bacteria
  2. Mycorrhizae
  3. Nodule-Forming And Other Nematodes
  4. Nitrogen Fixation
- D. Oxygen
  1. Transpiration
- E. Detoxification Effects Of
  7. Iron
  8. Aluminum
  9. Cadmium
  10. Mercury
  11. Salt
  12. Other Soil Constituents
- F. Pathogens



1. Chemical Repellent
  - (a) Glucosinolates (GSL), Which Release Pathogen-Controlling Isothiocyanates

G. Changes In Soil

1. Ph
2. Mineral Excess And Depletion
3. Rhizosheath

III. TRANSPORT OF MATERIALS IN PLANTS

Introduction: Uptake of the nutrients by the root and root hairs contributes a source-sink effect in a plant. The greater source of nutrients, the more sinks, such as stems, leaves, flowers, seeds, fruits, etc. can draw sustenance to grow. Thus, root hair development genes and gene products are useful to modulate the vigor and yield of the plant overall as well as of distinct cells, organs, or tissues of a plant. Examples are as follows:

A. Vigor

1. Plant Nutrition
2. Growth Rate
  - (a) Whole Plant, Including Height, Flowering Time, Etc.
  - (b) Seedling
  - (c) Coleoptile Elongation
  - (d) Young Leaves
  - (e) Stems
  - (f) Flowers
  - (g) Seeds
  - (h) Fruit
3. Yield
  - (a) Biomass
    - Fresh And Dry Weight During Any Time In Plant Life, Including Maturation And Senescence
  - (b) Number Of Flowers
  - (c) Number Of Seeds

- (d) Seed Yield
- (e) Number, Size, Weight, Harvest Index
  - Content And Composition, E.G. Amino Acid, Jasmonate, Oil, Protein And Starch
- (f) Fruit Yield
  - Number, Size, Weight, Harvest Index, Post Harvest Quality

#### Additional Uses of Plants with Modified Root Hairs

Plants with root hairs modified in one or more of the properties described above in I, II or III are used to provide:

- A. Higher vigor and yield of plant and harvested products due to pathogen resistance from conditioning the soil with plant-derived chemicals and/or more tolerance to stresses such as drought, flooding and anoxia
- B. Better Animal (Including Human) Nutrition
- C. Improved Dietary Mineral Nutrition
- D. Increased Plant Survival By Decreasing Lodging
- E. Better Plant Survival By:
  - (a) Decreased Lodging
  - (b) More Efficient Transport
  - (c) More Efficient Physiology
  - (d) More Efficient Metabolism
- F. Increased Yield Of Valuable Molecules

#### Root Hair Modulation

To regulate any of the phenotype(s) above, activities of one or more of the root hair genes or gene products is modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels are altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired

phenotype or metabolically and/or functionally assayed according to Dolan et al. (1993, Development 119: 71-84), Dolan et al. (1997, Development 124: 1789-98), Crawford and Glass (1998, Trends Plant Science 3: 389-95), Wang et al. (1998, PNAS USA 95: 15134-39), Gaxiola et al. (1998, PNAS USA 95: 4046-50), Apse et al. (1999, Science 285: 1256-58), Fisher and Long (1992, Nature 357: 655-60), Schneider et al. (1998, Genes Devel 12: 2013-21) and Hirsch (1999, Curr Opin Plant Biol. 2: 320-326).

**A.2.a. USE OF ROOT HAIR DEVELOPMENT GENES TO  
MODULATE BIOCHEMICAL ACTIVITIES**

The activities of one or more of the root hair development genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

Process	Biochemical Or Metabolic Activities And/Or Pathways	Citations Including Assays
Association Of Root Hair With Nitrogen Fixing Bacteria	<ul style="list-style-type: none"> <li>Functions Associated With Root Hair Curling And Signal Transduction</li> </ul>	Gage et al. (1996) J Bacteriol 178: 7159-66
Root Hair <ul style="list-style-type: none"> <li>Spacing</li> <li>Initiation</li> <li>Elongation</li> </ul>		Schneider et al. (1998) Genes Devel 12: 2013-21
Metabolism	<ul style="list-style-type: none"> <li>Organic Molecule Export</li> </ul>	Moody et al. (1988) Phytochemistry 27: 2857-61
	<ul style="list-style-type: none"> <li>Ion Export</li> </ul>	Uozumi et al. (2000) Plant Physiol 122: 1249-59 Frachisse et al. (2000) Plant J 21: 361-71
Nutrient Uptake	<ul style="list-style-type: none"> <li>Nutrient Uptake</li> </ul>	Frachisse et al. (2000) Plant J 21: 361-71 Uozumio et al. (2000) Plant Physiol 122: 1249-59

Other biological activities that can be modulated by the root hair genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

A.2.a. USE OF ROOT HAIR GENES, GENE COMPONENTS AND  
PRODUCT TO MODULATE TRANSCRIPTION LEVELS

Many genes are “up regulated” or “down regulated” in root hairs or associated with root hair formation because genes are regulated in networks. Thus some root hairs genes are useful to regulate the activities of many other genes, directly or indirectly to influence complex phenotypes. Examples of transcription profiles of root genes are described in the Table below with associated biological activities. “Up regulated” profiles are those where the mRNA levels are higher when the rhl gene is inhibited as compared to when rhl gene is not inhibited; and vice-versa for “down-regulated” profiles.

Transcript Levels	Type Of Genes	Physiological Consequences	Examples Of Biochemical Activity
Down Regulated Transcripts	<ul style="list-style-type: none"> <li>• Genes Expressed In Root Hair Development</li> <li>• Responders To Micro-Organismal Symbionts And Parasites</li> </ul>	<ul style="list-style-type: none"> <li>• Root Hair Formation</li> <li>• Microorganism Perception</li> <li>• Entrapment Of Microorganismal Symbionts</li> <li>• Nutrient Uptake</li> <li>• Synthesis Of Metabolites And/Or Proteins</li> <li>• Modulation Of Transduction Pathways</li> <li>• Specific Gene Transcription Initiation</li> <li>• Nutrient Uptake Enhancement</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic Enzymes</li> <li>• Change In Cell Membrane Structure And Potential</li> <li>• Kinases, Phosphatases, G-Proteins</li> <li>• Transcription Activators</li> <li>• Change In Chromatin Structure And/Or Localized DNA Topology</li> <li>• Cell Wall Proteins</li> </ul>
Up-Regulated Transcripts	<ul style="list-style-type: none"> <li>• Genes Repressed In Roots Making Hairs</li> <li>• Responders To Micro-Organismal Symbionts And</li> </ul>	<ul style="list-style-type: none"> <li>• Negative Regulation Of Hair Production Released</li> <li>• Changes In Pathways And Processes Operating In Cells</li> <li>• Changes In</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Kinases, Phosphatases, G-Proteins</li> <li>• Change In Chromatin Structure And/Or DNA Topology</li> <li>• Stability Of Factors</li> </ul>

Transcript Levels	Type Of Genes	Physiological Consequences	Examples Of Biochemical Activity
	Parasites  <ul style="list-style-type: none"> <li>Genes With Discontinued Expression Or UnsTable mRNAIn Presence Of Root Hairs And/Or Micro-Organismal Symbionts</li> </ul>	Metabolism	For Protein Synthesis And Degradation <ul style="list-style-type: none"> <li>Metabolic Enzymes</li> <li>Cell Wall Proteins</li> </ul>

Changes in the patterning or development of root hairs are the result of modulation of the activities of one or more of these many root hair genes and gene products. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield, especially when plants are growing in the presence of biotic or abiotic stresses or when they are growing in barren conditions or in soils depleted of certain minerals.

Root hair genes and gene products can act alone or in combination as described in the introduction. Of particular interest are combination of these genes and gene products with those that modulate stress tolerance and/or metabolism. Stress tolerance and metabolism genes and gene products are described in more detail in the sections below.

#### Use of Promoters of Root Hair Genes

Promoters of root hair development genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by root hair development or any of the following phenotypes or biological activities above. For example, any desired sequence can be

transcribed in similar temporal, tissue, or environmentally-specific patterns as the root hair genes when the desired sequence is operably linked to a promoter of a root hair responsive gene.

I.B. LEAF GENES, GENE COMPONENTS AND PRODUCTS

I.B.1. LEAF GENES, GENE COMPONENTS AND PRODUCTS

Leaves are responsible for producing most of the fixed carbon in a plant and are critical to plant productivity and survival. Great variability in leaf shapes and sizes is observed in nature. Leaves also exhibit varying degrees of complexity, ranging from simple to multi-compound. Leaf genes as defined here, not only modulate morphology, but also influence the shoot apical meristem, thereby affecting leaf arrangement on the shoot, internodes, nodes, axillary buds, photosynthetic capacity, carbon fixation, photorespiration and starch synthesis. Leaf genes elucidated here can be used to modify a number of traits of economic interest from leaf shape to plant yield, including stress tolerance, and to modify the efficiency of synthesis and accumulation of specific metabolites and macromolecules.

B.1.a. IDENTIFICATION OF LEAF GENE, GENE COMPONENTS  
AND PRODUCTS

Leaf genes identified herein are defined as genes, active or potentially active to greater extent in leaves than in some other organs of the plant or as genes that affect leaf properties. These genes and gene components are useful for modulating one or more processes in or functions of leaves, as described below, to improve plant traits ranging from yield to stress tolerance. Examples of such leaf genes and gene products are shown in the Reference and Sequence Tables and sequences encoding polypeptides of the Protein Group and Protein Group Matrix tables or fragments thereof, Knock-In, Knock-Out and MA\_diff and/or AFLP\_diff Tables. The biochemical functions of the protein products of many of these genes determined from comparisons with known proteins are also given in the Reference tables.

Leaf Genes Identified by Phenotypic Observations

Some leaf genes were discovered and characterized from a much larger set of genes by experiments designed to find genes that cause phenotypic changes in leaf, petiole, internode, and cotyledon morphology.

In these experiments, leaf genes were identified by either (1) ectopic expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then cultivated and



one or more of the following leaf phenotypes, which varied from the parental "wild-type", were observed:

A. Changes In Seedling Stage Cotyledons

- Cup Shaped
- Curled
- Horizontally Oblong
- Long Petioles
- Short Petioles
- Silver
- Tricot
- Wilted

B. Changes In Rosette And Flowering Stage Leaf Shapes

- Cordate
- Cup-Shaped
- Curled
- Fused
- Lanceolate
- Lobed
- Long Petioles
- Short Petioles
- Oval
- Ovate
- Serrate
- Trident
- Undulate
- Vertically Oblong

C. Changes In Cauline, Flowering Leaf Shape

- Misshapen
- Other

D. Changes In Leaf Pigment

- Albino
- Dark Green Pigment
- High Anthocyanin
- Interveinal Chlorosis
- Yellow Pigment

E. Changes In Leaf Size

F. Changes In Seedling Stage Hypocotyl

- Long
- Short

G. Changes In Leaf Number

H. Changes In Wax Deposition

- Glossy Rosette And Flowering Stage Leaves
- Altered Wax Deposition On The Bolt

#### Leaf Genes Identified by Differential Expression

Also, leaf genes were identified in experiments in which the concentration of mRNA products in the leaf, or stem, or Knock-out mutant 3642-1 were compared with to a control. The MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108477, 108512, 108497, 108498, 108598). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Leaf genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Leaf Genes Identified By Cluster Analyses Of Differential Expression

##### Leaf Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table

indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Leaf genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108477, 108512, 108497, 108498, 108598 of the MA\_diff and/or AFLP\_diff table(s).

#### Leaf Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Leaf genes. A group in the MA\_clust is considered a Leaf pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Leaf Genes Identified By Amino Acid Sequence Similarity

Leaf genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Leaf genes. Groups of Leaf genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Leaf pathway or network is a group of proteins that also exhibits Leaf functions/utilities.

It is assumed that (i) the genes preferentially expressed in leaves are concerned with specifying leaf structures and the synthesis of all the constituent molecules and (ii) that the genes repressed in leaves specify products that are not required in leaves or that could inhibit normal leaf development and function.

Examples of phenotypes, biochemical activities, and transcription profiles that are modulated by using selected members of these genes and gene products, singly or in combination, are described below.

B.1.b. USE OF LEAF GENES, GENES COMPONENTS AND  
PRODUCTS TO MODULATE PHENOTYPES

Leaves are critical for the performance and industrial utility of plants. There is extensive evidence that the number, size, shape, position, timing of synthesis, timing of senescence and chemical constitution are very important for agriculture, horticulture and uses of plants as chemical factories for making valuable molecules. Many improvements already demonstrated over past decades have involved genetic modifications to leaves. Therefore, the leaf genes and gene components of this invention offer considerable opportunities for further improving plants for industrial purposes. When the leaf genes and/or gene components are mutated or regulated differently, they are capable of modulating one or more of the processes determining leaf structure and/or function including (I) development; (II) interaction with the environment and (III) photosynthesis and metabolism.

I. DEVELOPMENT

The leaf genes, gene components and products of the instant invention are useful to modulate one or more processes of the stages of leaf morphogenesis including: stage 1- organogenesis that gives rise to the leaf primordium; stage 2- delimiting basic morphological domains; and stage 3- a coordinated processes of cell division, expansion, and differentiation. Leaf genes include those genes that terminate as well as initiate leaf development. Modulating any or all of the processes leads to beneficial effects either at specific locations or throughout the plant.

A. Gene Sequences Affecting Types of Leaves

Applicants provide with these genes, gene components and gene products the means to modulate one or more of the following types of leaves, shoots, and stems:

1. Cotyledons
2. Major Leaves
3. Cauline Leaves
4. Petioles

B. Gene Sequences Affecting Cell properties

Leaf genes, gene components and gene products are useful to modulate changes in:

1. Cell Size
2. Cell Division, Rate And Direction
3. Cell Elongation
4. Cell Differentiation
5. Stomata Size, Number, Spacing And Activity
6. Trichome Size And Number (For More Details See Section On Trichome Genes)
7. Xylem And Phloem Cell Numbers
8. Cell Wall Composition
9. All Cell Types

C. Gene Sequences Affecting Leaf Architecture:

The following properties of a leaf are useful to modulate to change overall leaf architecture:

1. Veination – Improvements in photosynthetic efficiency, stress tolerance efficiency of solute and nutrient movement to and from the leaf are accomplished by increases or decreases in:
  - (a) Vein Placement
  - (b) Number Of Cells In The Vein
2. Shape
  - (a) Elongated Versus Rounded
  - (b) Symmetry, Around Either
    - Abaxial-Adaxial (Dorsiventral) Axis
    - Apical-Basal (Proximodistal) Axis
    - Margin-Blade-Midrib (Lateral) Axis
3. Branching – Improved plant performance to biotic and abiotic stress in heavy density planting is achieved by increases or decreases in:
  - (a) Leaf branch position
  - (b) Leaf branch length

### G. Genes Sequences Influencing Leaf Responses

Shoot apical meristem cells differentiate to become leaf primordia that eventually develop into leaves. The genes, gene components and gene products of this invention are useful to modulate any one or all of these growth and development processes, by affecting timing and rate or planes of cell divisions for example, in response to the internal plant stimuli and/or programs listed below:

1. Embryogenesis
2. Germination
3. Hormones
  - (a) Auxin (For More Details See The Section On Auxin-Responsive Genes)
4. Leaf Senescence
5. Phototropism
6. Coordination Of Leaf Growth And Development With That Of Other Organs
  - (a) Roots
  - (b) Flowers
  - (c) Seeds
  - (d) Fruits
  - (e) Stems
7. Stress-Related Programs

## II. INTERACTION WITH THE ENVIRONMENT

Leaves are the main sites of photosynthesis and have various adaptations for that purpose. Flat laminae provide a large surface for absorbing sunlight; leaves are rich in chloroplasts and mitochondria; stomata in the lower surface of the laminae allow gases to pass into and out of the leaves including water; and an extensive network of veins brings water and minerals into the leaves and transports the sugar products produced by photosynthesis to the rest of the plant. Examples of leaf properties or activities that are modulated by leaf genes, gene components and their products to facilitate interactions between a plant and the environment include:

- A. Pigment Accumulation (See The Section On Viability Genes For More Detail)
- B. Wax accumulation on the surface of leaves – Improved protection of young leaves from water borne pathogen attack such as Downey Mildew with increased wax production.
- C. Oxygen Gain/Loss Control
- D. Carbon Dioxide Gain/Loss Control
- E. Water Gain/Loss Control
- F. Nutrient Transport
- G. Light Harvesting
- H. Chloroplast Biogenesis
- I. Circadian Rhythm Control
- J. Light/Dark Adaptation
- K. Defense Systems Against Biotic And Abiotic Stresses
- L. Metabolite Accumulation
- M. Secondary metabolite production in leaf mesophyll, epidermis and trichomes.
  - (a) Increases in antifeeding secondary metabolites such as strictosiden reduce herbivory.
  - (b) Decreases in secondary metabolites improve plants as forage by reducing allergens or undigestible compounds.

### III. PHOTOSYNTHESIS AND METABOLISM

Many of the uses for plants depend on the success of leaves as the powerhouses for plant growth, their ability to withstand stresses and their chemical composition. Leaves are organs with many different cell types and structures. Most genes of a plant are active in leaves and therefore leaves have very diverse of pathways and physiological processes. Examples of such pathways and processes that are modulated by leaf genes, gene components and products include:

- A. Photosynthesis
- B. Sugar Metabolism

- C. Starch Synthesis
- D. Starch Degradation
- E. Nitrate And Ammonia Metabolism
- F. Amino Acid Biosynthesis, Transport
- G. Protein Biosynthesis
- H. DNA Replication, Repair
- I. Lipid Biosynthesis And Breakdown
- J. Protein Biosynthesis, Storage And Breakdown
- K. Nucleotide Transport And Metabolism
- L. Cell Envelope Biogenesis
- M. Membrane Formation
- N. Mitochondrial And Chloroplast Biogenesis
- O. Transcription And RNA Metabolism
- P. Vitamin Biosynthesis
- Q. Steroid And Terpenoid Biosynthesis
- R. Devise Secondary Metabolite Synthesis
- S. Co-Enzyme Metabolism
- T. Flavonoid Biosynthesis And Degradation
- U. Synthesis Of Waxes
- V. Glyoxylate Metabolism
- W. Hormone Perception And Response Pathways

#### Uses of Plants that Are Modified as Described Above

Altering leaf genes or gene products in a plant modifies one or more of the following plant traits, to make the plants more useful for specific purposes in agriculture, horticulture and for the production of valuable molecules. The useful plants have at least one of the following:

- A. A higher yield of leaves and their molecular constituents due to different:
  - 1. Number, Size, Weight, Harvest Index
  - 2. Composition Including And Amounts And Types Of Carbohydrates, Proteins, Oils, Waxes, Etc.



3.      Photosynthetic Efficiency E.G. Reduced Photorespiration
  4.      Absorption Of Water And Nutrients To Enhance Yields,  
         Including Under Stresses Such As High Light, Herbicides,  
         And Heat.
  5.      Pathways To Accumulate New Valuable Molecules.
- B.      More optimal leaf shape and architecture – enhancing photosynthesis and  
         enhancing appeal in ornamental species
- (a)     Size
  - (b)     Number
  - (c)     Pigment
  - (d)     Aroma
- C.      A better overall plant architecture – enhancing photosynthesis and  
         enhancing appeal in ornamental species
- (a)     Petals
  - (b)     Sepals
  - (c)     Stamens
  - (d)     Carpels
- D.      Better shade avoidance for maximizing photosynthesis by, for example,  
         altering leaf placement, to improve light capture and photosynthetic  
         efficiency, thereby increasing yields
- E.      Reduced negative effects of high planting density, by altering leaf  
         placement to be more vertical instead of parallel to the ground, for  
         instance
- F.      More resistance to the deleterious effects of wind and mechanical damage.
- G.      Better stress tolerance, including without limitation
1.      Drought resistance, by decreasing water loss, for example
  2.      Pathogen resistance, including, for instance,  
         Insect resistance through internal insecticide levels and  
         optimizing the leaf shape to prevent runoff of insecticides
- H.      Better overall yield and vigor

Plant yield of biomass and of constituent molecules and plant vigor are modulated to create benefits by genetically changing:

1. Growth Rate Of
  - (a) Whole Plant, Including Height, Flowering Time, Etc.
  - (b) Seedling
  - (c) Coleoptile Elongation
  - (d) Young Leaves
  - (e) Flowers
  - (f) Seeds
  - (g) Fruit
2. Biomass
  - (a) Fresh And Dry Weight During Any Time In Plant Life, Including Maturation And Senescence
  - (b) Number Of Flowers
  - (c) Seed Yield Including, For Example,
    - Number, Size, Weight, Harvest Index
    - Content And Composition, E.G. Amino Acid, Jasmonate, Oil, Protein And Starch
  - (d) Fruit Yield
    - Number, Size, Weight, Harvest Index
    - Content And Composition, E.G. Amino Acid, Jasmonate, Oil, Protein And Starch

To change any of the phenotype(s) in I, II, or III above, activities of one or more of the leaf genes or gene products are modulated in an organism and the consequence evaluated by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels are altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (Methods. Mol. Biol. 82:259-266 (1998)) with leaf gene constructs and/or screened for variants as in Winkler et al., Plant Physiol. 118: 743-50 (1998) and visually inspected for the desired phenotype and metabolically and/or functionally assayed for altered levels of relevant molecules.

B.1.c. USE OF LEAF GENES, GENE COMPONENTS AND PRODUCTS TO  
MODULATE BIOCHEMICAL ACTIVITIES

Leaves are complex organs and their structure, function and properties result from the integration of many processes and biochemical activities. Some of these are known from the published literature and some can be deduced from the genes and their products described in this application. Leaf genes, and gene components are used singly or in combination to modify these processes and biochemical activities and hence modify the phenotypic and trait characteristics described above. Examples of the processes and metabolic activities are given in the Table below. The resulting changes are measured according to the citations included in the Table.

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Metabolism – anabolic and catabolic	<ul style="list-style-type: none"> <li>• Farnesylation</li> <li>• Cell Wall Biosynthesis</li> <li>• Nitrogen Metabolism</li> <li>• Secondary Metabolite Biosynthesis and Degradation</li> </ul>	<p>Pei et al., <u>Science</u> <u>282</u>: 287-290 (1998); Cutler et al., <u>Science</u> <u>273</u>: 1239 (1996)</p> <p>Goupil et al., <u>J Exptl. Botany</u> <u>49</u>:1855-62 (1998)</p> <p>Walch-Liu et al., <u>J Exppt. Botany</u> <u>51</u>, 227-237 (2000)</p>
Water Conservation And Resistance To Drought And Other Related Stresses	<ul style="list-style-type: none"> <li>• Stomatal Development And Physiology</li> <li>• Production of polyols</li> <li>• Regulation of salt concentration</li> <li>• ABA response(s)</li> <li>• Ca<sup>2+</sup> Accumulation</li> <li>• K<sup>+</sup> Fluxes</li> </ul>	<p>Allen et al., <u>Plant Cell</u> <u>11</u>: 1785-1798 (1999)</p> <p>Li et al., <u>Science</u> <u>287</u>: 300-303 (2000)</p> <p>Burnett et al., <u>J Exptl. Botany</u> <u>51</u>: 197-205 (2000)</p> <p>Raschke, In: <u>Stomatal Function</u>, Zeiger et al. Eds., 253-279 (1987)</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Transport Anion and Cation Fluxes	<ul style="list-style-type: none"> <li>• Na<sup>+</sup> Fluxes</li> <li>• Receptor – ligand binding</li> <li>• Anion and Cation fluxes</li> </ul>	<p>Lacombe et al., <u>Plant Cell</u> <b>12</b>: 837-51 (2000);</p> <p>Wang et al., <u>Plant Physiol.</u> <b>118</b>:1421-1429 (1998);</p> <p>Shi et al., <u>Plant Cell</u> <b>11</b>: 2393-2406 (1999)</p> <p>Gaymard et al., <u>Cell</u> <b>94</b>:647-655 (1998)</p> <p>Jonak et al., <u>Proc. Natl. Acad. Sci.</u> <b>93</b>: 11274-79 (1996);</p> <p>Sheen, <u>Proc. Natl. Acad. Sci.</u> <b>95</b>: 975-80 (1998);</p> <p>Allen et al., <u>Plant Cell</u> <b>11</b>: 1785-98 (1999)</p>
Carbon Fixation	<ul style="list-style-type: none"> <li>• Calvin Cycle <ul style="list-style-type: none"> <li>- Photorespiration</li> <li>- Oxygen evolution</li> <li>- RuBisCO</li> </ul> </li> <li>• Chlorophyll metabolism</li> <li>• Chloroplast Biogenesis and Metabolism</li> <li>• Fatty Acid and Lipid Biosynthesis</li> <li>• Glyoxylate metabolism</li> <li>• Sugar Transport</li> <li>• Starch Biosynthesis and</li> </ul>	<p>Wingler et al., <u>Philo Trans R Soe Lond B Biol Sci</u> <b>355</b>, 1517-1529 (2000);</p> <p>Palecanda et al., <u>Plant Mol Biol</u> <b>46</b>, 89-97 (2001);</p> <p>Baker et al., <u>J Exp Bot</u> <b>52</b>, 615-621 (2001)</p> <p>Chen et al., <u>Acta Biochim Pol</u> <b>41</b>, 447-457 (1999)</p> <p>Imlau et al., <u>PlantCell</u> <b>11</b>, 309-322 (1999)</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Degradation	
Hormone Perception and Growth	<ul style="list-style-type: none"> <li>Hormone Receptors and Downstream Pathways for                             <ul style="list-style-type: none"> <li>ethylene</li> <li>jasmonic acid</li> <li>brassinosteroid</li> <li>gibberellin</li> <li>auxin</li> <li>cytokinin</li> </ul> </li> <li>Activation Of Specific Kinases And Phosphatases</li> </ul>	<p>Tieman et al., <u>Plant J</u> <u>26</u>, 47-58 (2001)</p> <p>Hilpert et al., <u>Plant J</u> <u>26</u>, 435-446 (2001)</p> <p>Wenzel et al., <u>Plant Phys</u> <u>124</u>, 813-822 (2000)</p> <p>Dengler and Kang, <u>Curr Opin Plant Biol</u> <u>4</u>, 50-56 (2001)</p> <p>Tantikanjana et al., <u>Genes Dev</u> <u>15</u>, 1577-1580 (2001)</p>

Other biological activities that are modulated by the leaf genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table, for example.

B.1.d. USE OF LEAF GENES, GENE COMPONENTS AND PRODUCTS TO  
MODULATE TRANSCRIPTION LEVELS

5 The expression of many genes is “upregulated” or downregulated” in leaves because  
some leaf genes and their products are integrated into complex networks that regulate  
transcription of many other genes. Some leaf genes, gene components and products are therefore  
useful for modifying the transcription of other genes and hence complex phenotypes, as  
described above. Profiles of leaf gene activities are described in the Table below with associated  
biological activities. “Up-regulated” profiles are those where the mRNA transcript levels are  
10 higher in leaves as compared to the plant as a whole. “Down-regulated” profiles represent higher  
transcript levels in the whole plant as compared to leaf tissue only.

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS WITH MODIFIED LEVELS
Up Regulated Transcripts	Genes Involved In Leaf Cell Differentiation, Cell Division, Cell Expansion	<ul style="list-style-type: none"> <li>• Leaf Cells Proliferate And Differentiate;</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors, Signal Transduction Proteins, Kinase And Phosphatases</li> </ul>
	Genes Involved In Positive Regulation Of Leaf Genes	<ul style="list-style-type: none"> <li>• Leaf Structures Form And Expand</li> </ul>	<ul style="list-style-type: none"> <li>• Chromatin Remodeling</li> <li>• Hormone Biosynthesis Enzymes</li> <li>• Receptors</li> </ul>
	Repressors Of Root And Other Non Leaf Cell Types	<ul style="list-style-type: none"> <li>• Photosynthesis And Plastid Differentiation</li> </ul>	<ul style="list-style-type: none"> <li>• Light Harvesting Coupled To ATP Production</li> <li>• Chlorophyll Biosynthesis</li> </ul>
	Genes Involved In Photosynthesis		

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS WITH MODIFIED LEVELS
		<ul style="list-style-type: none"> <li>• Calvin Cycle Activated</li> <li>• Chloroplast Biogenesis And Plastid Differentiation Activated</li> </ul>	<ul style="list-style-type: none"> <li>• Ribulose Biphosphate Carboxylase</li> <li>• Chloroplast Membranes Synthesis</li> <li>• Chloroplast Ribosome Biogenesis</li> </ul>
	Other Genes Involved In Metabolism	<ul style="list-style-type: none"> <li>• Starch Biosynthesis</li> <li>• Lipid Biosynthesis</li> <li>• Nitrogen Metabolism – NO<sub>3</sub> Reduced And Amino Acids Made</li> <li>• Secondary Metabolites Produced</li> </ul>	<ul style="list-style-type: none"> <li>• Starch Synthase</li> <li>• Nitrate Reductase</li> <li>• Terpenoid Biosynthesis</li> <li>• Transcription Factors</li> <li>• Transporters</li> <li>• Kinases</li> <li>• Phosphatases And Signal Transduction Protein</li> <li>• Chromatin Structure Modulators</li> </ul>



TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS WITH MODIFIED LEVELS
Down Regulated Genes	Genes Involved In Negative Regulation Of Leaf Genes	<ul style="list-style-type: none"> <li>• Leaf Genes Activated And Leaf Functions Induced;</li> <li>• Dark-Adapted Metabolism Suppressed</li> <li>• Meristematic Genes Suppressed</li> <li>• Leaf Metabolic Pathways Induced</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Signal Transduction Proteins – Kinases And Phosphatases</li> <li>• Metabolic Enzymes</li> <li>• Chromatin Remodeling Proteins</li> </ul>

While leaf polynucleotides and gene products are used singly, combinations of these polynucleotides are often better to optimize new growth and development patterns. Useful combinations include different leaf polynucleotides and/or gene products with a hormone responsive polynucleotide. These combinations are useful because of the interactions that exist between hormone-regulated pathways, nutritional pathways and development.

#### Use of Leaf Gene Promoters

Promoters of leaf genes are useful for transcription of desired polynucleotides, both plant and non-plant. If the leaf gene is expressed only in leaves, or specifically in certain kinds of leaf cells, the promoter is used to drive the synthesis of proteins specifically in those cells. For example, extra copies of carbohydrate transporter cDNAs operably linked to a leaf gene promoter and inserted into a plant increase the “sink” strength of leaves. Similarly, leaf

promoters are used to drive transcription of metabolic enzymes that alter the oil, starch, protein, or fiber contents of a leaf. Alternatively, leaf promoters direct expression of non-plant genes that can, for instance, confer insect resistance specifically to a leaf. Additionally the promoters are used to synthesize an antisense mRNA copy of a gene to inactivate the normal gene expression into protein. The promoters are used to drive synthesis of sense RNAs to inactivate protein production via RNA interference.

### 1.B.2. TRICHOME GENES AND GENE COMPONENTS

Trichomes, defined as hair-like structures that extend from the epidermis of aerial tissues, are present on the surface of most terrestrial plants. Plant trichomes display a diverse set of structures, and many plants contain several types of trichomes on a single leaf. The presence of trichomes can increase the boundary layer thickness between the epidermal tissue and the environment, and can reduce heat and water loss. In many species, trichomes are thought to protect the plant against insect or pathogen attack, either by secreting chemical components or by physically limiting insect access to or mobility on vegetative tissues. The stellate trichomes of Arabidopsis do not have a secretory anatomy, but at a functional level, they might limit herbivore access to the leaf in the field. In addition, trichomes are known to secrete economically valuable substances, such as menthol in mint plants.

#### B.2.a. IDENTIFICATION OF TRICHOME GENES, GENE COMPONENTS AND PRODUCTS

Trichome genes identified herein are defined as genes or gene components capable of modulating one or more processes in or functions of a trichome, as described below. These genes, their components and products are useful for modulating diverse plant traits from production of secondary metabolites to pathogen resistance. Examples of such trichome genes and gene products are shown in the Reference and Sequence Tables and sequences encoding polypeptides of the Protein Group and Protein Group Matrix tables or fragments thereof, Knock-in, Knock-out, MA-diff and MA-clust. The biochemical functions of the protein products of many of these genes determined from comparisons with known proteins are also given in the Reference tables.

#### Trichome Genes Identified by Phenotypic Observation

Trichome genes were discovered and characterized from a much larger set of genes by experiments designed to find genes that cause phenotypic changes in trichome number and morphology on leaf, internode, cotyledon, petiole, and inflorescence. In these experiments, trichome genes were identified by either (1) ectopic expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then cultivated and one or more of the following phenotypes, which varied from parental "wild-type", were observed: (1) trichome

number; (2) trichome spacing (clustering); or (3) trichome branching. The genes regulating trichome phenotypes are identified in the Knock-In and Knock-Out Tables.

#### Trichome Genes Identified by Differential Expression

Trichome genes were also discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA products are associated specifically or preferentially with trichomes. These experiments made use of an Arabidopsis glabrous mutant and a hairy mutant. By comparing gene expression profiles of the glabrous mutant with those of the hairy mutant grown under identical conditions, genes specifically or preferentially expressed in trichomes were revealed. The MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108452). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Trichome genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Trichome Genes Identified By Cluster Analyses Of Differential Expression

##### Trichome Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Trichome genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108452 of the MA\_diff and/or AFLP\_diff table(s).

##### Trichome Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Trichome genes. A group in the MA\_clust is considered a Trichome pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Trichome Genes Identified By Amino Acid Sequence Similarity

Trichome genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Trichome genes. Groups of Trichome genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Trichome pathway or network is a group of proteins that also exhibits Trichome functions/utilities.

It is assumed that the genes differentially expressed in trichomes or leaves producing trichomes are concerned with specifying trichomes and their functions and therefore modulations of such genes and their products modify trichomes and their products.

Examples of phenotypes, biochemical activities, and transcription profiles that can be modulated by selected numbers of these genes and gene products singly or in combinations are described above and below.

#### B.2.b. USE OF TRICHOME GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE PHENOTYPES

Trichome genes of the instant invention, when mutated or activated differently, are useful for modulating one or more processes of trichome structure and/or function including: (I) development; (II) plant stress tolerance; and (III) biosynthesis or secretion of trichome-specific molecules. Trichome genes, components and gene products are useful to alter or modulate one or more of the following phenotypes:

##### I. Development

Trichome differentiation is integrated with leaf development, hormone levels and the vegetative development phase. The first trichome at the leaf tip appears only after the leaf grows to ~100  $\mu\text{m}$  in length. Subsequent events proceed basipetally as the leaf grows. As leaf development progresses, cell division patterns become less regular and islands of dividing cells

can be observed among differentiated pavement cells with their characteristic lobed morphology. Trichome initiation in the expanding leaf occurs within these islands of cells and often defines points along the perimeter of a circle, with an existing trichome defining the center.

Once a cell enters the trichome pathway it undergoes an elaborate morphogenesis program that has been divided into different stages based on specific morphological hallmarks.

Selected members of the genes, gene components and gene products of this invention are useful to modulate any one or all of these growth and development processes by affecting rate, timing, direction and size, for example. Trichome genes can also affect trichome number and the organs on which they occur. The following can be modulated by these genes:

A. Types of Trichomes

Applicants provide the means to modulate one or more of the following types of trichomes with the genes and gene products of this invention:

1. Glandular trichomes
2. Stellate trichomes

B. Cell properties

Trichome genes and gene products are useful to modulate cellular changes in:

1. Cell size
2. Cell division rate and direction
3. Cell elongation
4. Cell differentiation
5. Secretory cells
6. Trichome number (Average trichome number per leaf for mint:13,500,000).
7. Cell walls
8. Cell death
9. Response to reactive oxygen species

C. Trichome Architecture

The following parts or arrangement of trichomes on leaves can be modulated to affect trichome and/or leaf architecture:

1. Trichome cell structure
2. Placement on leaf
3. Secretory systems

D. Trichome responses

Selected members of trichome genes, gene components and gene products of this invention are useful to modulate any one or all of the growth and development processes above; as in timing and rate, for example. In addition, the polynucleotides and polypeptides of the invention can control the response of these processes to internal plant programs and signaling molecules such as:

1. Leaf development
2. Hormones:
  - (a) abscisic acid (for more details see the section on abscisic acid-responsive genes)
  - (b) auxin (for more details see the section on auxin-responsive genes)
  - (c) cytokinin (for more details see the section on cytokinin-responsive genes)
  - (d) gibberellins (for more details see the section on gibberellin-responsive genes)
  - (e) brassinosteroids (for more details see the section on brassinosteroid-responsive genes)
3. Apoptosis
4. Coordinated trichome growth and development in:
  - (a) Flowers
  - (b) Stems
  - (c) Petioles
  - (d) Cotyledons
  - (e) Hypocotyls

## II. Stress Tolerance

The physical characteristics of trichomes as well as the substances secreted by trichomes are useful in protecting the plant from both biotic and abiotic attacks. Thus, selected trichome genes and gene products can be used to help protect distinct cells, organs, or tissues as well as overall plant yield and vigor. Examples of stresses, tolerances to which are modulated by trichome genes and gene products, are as follows:

- A. Drought, e.g., trichome number variation can decrease the surface area that allows evaporation
- B. Heat, e.g., trichomes can produce shade and provide protection for meristems
- C. Salt
- D. Insects, e.g., trichomes can prevent insects from settling on plant surfaces
- E. Herbivory, e.g., trichomes can produce harmful chemicals
- F. Ultraviolet light

## III. Biosynthesis, Accumulation or Secretion of Metabolites

The glandular trichomes from various species are shown to secrete and, sometimes, locally synthesize a number of substances including salt, monoterpenes and sesquiterpenes, terpenoids, exudate, insect entrapping substances, antifeedants, pheromones, and others. Therefore, trichome genes can be used to modulate the synthesis, accumulation and secretion of a large number of metabolites especially related to trichome biology. Some are synthesized in response to biotic and abiotic stresses. For a more detailed description of these metabolites see the section "Use of Trichome Genes to Modulate Biochemical Activities" below.

## Uses of Plants that Are Modified as Described Above

Altering trichome properties is useful for modifying one or more of the following plant traits making the plants more useful in agriculture, horticulture and for the production of valuable molecules.



- A. Production of specific carbohydrates, proteins, oils, aromas, flavors, pigments, secondary metabolites such as menthol (and other monoterpenes), etc., that can be used in situ or purified and used in a wide variety of industries.
- B. Increased production of molecules synthesized in trichomes by increasing the trichome number on different plant organs, such as cotyledons, leaves, hypocotyls, stems, petioles, etc.
- C. Increased cotton fibers per boll due to decreased numbers of trichomes that reduces insect hiding and contamination
- D. More optimal growth rate of a whole plant or specific parts of a plant due to more optimal trichome cellular development and the better resistance to biotic /abiotic stresses. Examples of specific plant parts include, without limitation,
1. Whole plant
  2. Seedling
  3. Coleoptile elongation
  4. Young leaves
  5. Flowers
  6. Seeds
  7. Fruit
- E. Increased harvested yield of plants, organs and their constituent molecules
1. Biomass
    - (a) Fresh and dry weight during any time in plant life, including maturation and senescence
    - (b) Number of flowers
    - (c) Seed yield
      - Number, size, weight, harvest index
      - Content and composition, e.g. amino acid, jasmonate, oil, protein and starch
    - (d) Fruit yield
      - Number, size, weight, harvest index, post harvest quality
      - Content and composition, e.g. amino acid, jasmonate, oil, protein and starch

To regulate any of the phenotype(s) above, activities of one or more of the trichome genes or gene products can be modulated in an organism and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the  
5 procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (Methods, Mol. Biol. 82:259-266 (1998)) and/or screened for variants as in Winkler et al., Plant Physiol. 118: 743-50 (1998) and visually inspected for the desired phenotype or metabolically and/or functionally assayed.

B.2.c. USE OF TRICHOME GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The phenotype traits outlined above result from the integration of many cellular trichome associated processes and biochemical activities. Some of these are known from published literature and some can be deduced from the genes discovered in the MA Tables, etc. One or more of these trichome genes, gene components and products are useful to modulate these cellular processes, biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation And Development	<ul style="list-style-type: none"> <li>Cell wall biosynthetic enzymes</li> <li>Cell fate determination proteins</li> <li>Major pathways of carbon and nitrogen metabolism</li> </ul>	<p>Molhoj et. al. (2001). Plant Mol.Biol. 46, 263-275</p> <p>Krishnakumar and Oppenheimer (1999). Development 1221, 3079-3088.</p> <p>Kroumova et al. (1994). PNAS 91, 11437-11441</p>
Water Conservation And Resistance To Drought And Other Related Stresses	<ul style="list-style-type: none"> <li>Cytoskeleton and Trichome morphology and spacing controls</li> </ul>	<p>Schnittger et al. (1999). Plant Cell 11, 1105-1116</p> <p>Hulskamp et al (1994). Cell 76, 555-566</p>
Trichome exudate	<ul style="list-style-type: none"> <li>Insect repellant</li> </ul>	Insects and The Plant Surface, pp 151-172,

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		Edward Arnold, London (1986)
Terpenoid biosynthesis including monoterpenes and sesquiterpenes	<ul style="list-style-type: none"> <li>• Terpenoid biosynthesis enzymes including: <ul style="list-style-type: none"> <li>• Farnesyltransferase</li> <li>• Geranylgeranyl-diphosphate synthase</li> <li>• Geranyltransferase</li> <li>• Farnesyl-diphosphate synthase</li> <li>• Dimethylallyltransferase</li> <li>• Geranyl-diphosphate synthase</li> </ul> </li> </ul>	Alonso et al. (1992). J. Biol. Chem. 267, 7582-7587 Rajonarivony et al (1992). Arch. Biochem. Biophys. 299, 77-82
H <sub>2</sub> O <sub>2</sub> accumulation and activation of SAR	<ul style="list-style-type: none"> <li>• NADPH oxidase (subunit) synthesis and function</li> </ul>	Alvarez et al (1998) Cell 92, 773-784 Grant Orozco-Cardenas and Ryan (1999) PNAS 96, 6553-6557
Antifeedants biosynthesis and secretion	<ul style="list-style-type: none"> <li>• Lactone biosynthesis enzymes</li> </ul>	Paruch et al. (2000). J. Agric. Food Chem. 48, 4973-4977
Pheromone biosynthesis and secretion	<ul style="list-style-type: none"> <li>• Farnesine biosynthesis enzymes</li> </ul>	Teal et al. (1999) Arch. Insect Biochem Physiol. 42, 225-232
Endoreplication	<ul style="list-style-type: none"> <li>• Cyclin and cyclin dependant</li> </ul>	De Veylder et al. (2001)

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	kinases	Plant Cell 13, 1653-1668 De Veylder et al. (2001) Plant J. 25, 617-626

Specific enzyme and other activities associated with the functions of individual trichome genes that can be modulated by the trichome genes and gene products are listed in the Reference tables where the functions of individual genes and their products are listed. Assays for detecting such biological activities are described in the Protein Domain table, for example.

B.2.d. USE OF TRICHOME GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE PHENOTYPES BY MODULATING  
TRANSCRIPTION LEVELS OF OTHER GENES

Many of the genes are “up regulated” or “down regulated” in trichomes because they are regulated as members of networks or cascade of genes under the control of regulatory genes. Thus some trichome genes are useful to influence levels of other genes and so orchestrate the complex phenotypes. Examples of the types of genes with altered transcript levels in trichomes are described in the Table below, together with associated biological activities. “Up-regulated” profiles are those where the mRNA levels are higher in the glabrous plants as compared to the “hairy” plant. “Down-regulated” profiles represent higher transcript levels in the “hairy” plant as compared to the glabrous plant.

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
Up Regulated Transcripts	Genes active in suppressing trichome formation	<ul style="list-style-type: none"> <li>• Changes in Hormone Perception</li> <li>• Changes in Hormone Biosynthesis</li> <li>• Changes in Specific Gene Transcription Initiation</li> <li>• Changes in cytoskeleton and cell wall assembly and structure</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Transporters</li> <li>• Change In Cell G-proteins</li> <li>• Kinases And Phosphatases</li> <li>• Transcription factors</li> <li>• Ca-binding proteins</li> <li>• Transcription Activators</li> <li>• Change In Chromatin Structure And/Or Localized DNA Topology</li> <li>• Specific Factors (Initiation And Elongation) For Protein Synthesis</li> <li>• Maintenance Of mRNA Stability</li> <li>• Maintenance Of</li> </ul>

131

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
	Genes associated with trichome-specific metabolic pathways	<ul style="list-style-type: none"> <li>• Changes in terpenoid biosynthesis</li> <li>• Changes in antifeedant and pheromone biosynthesis</li> </ul>	

While trichome polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth, development and leaf biochemistry. Combinations of trichome polynucleotide(s) and/or gene product(s) with genes or gene products involved in leaf development, hormone responses, or vegetative development are useful because trichome development is integrated with these processes.

#### Use of Promoters of Trichome Genes

Promoters of trichome genes are useful for facilitating transcription of desired polynucleotides, both plant and non-plant in trichomes. For example, extra copies of existing terpenoid synthesis coding sequences can be operably linked to a trichome gene promoter and inserted into a plant to increase the terpenoids in the trichome. Alternatively, trichome promoters can direct expression of non-plant genes or genes from another plant species that can, for instance, lead to new terpenoids being made. The promoters can also be operably linked to antisense copies of coding sequences to achieve down regulation of these gene products in cells.



PATENT  
Attorney Docket No. 2750-1481P

Attorney Docket No. 2750-1481P

I.B.3. CHLOROPLAST GENES, GENE COMPONENTS AND PRODUCTS

10 The chloroplast is a complex and specialized organelle in plant cells. Its complexity  
comes from the fact that it has at least six suborganellar compartments subdivided by double-  
5 membrane envelope and internal thylakoid membranes. It is specialized to carry out different  
biologically important processes including photosynthesis and amino acid and fatty acid  
biosynthesis. The biogenesis and development of chloroplast from its progenitor (the  
proplastid) and the conversion of one form of plastid to another (e.g., from chloroplast to  
amyloplast) depends on several factors that include the developmental and physiological states of  
the cells.

One of the contributing problems that complicate the biogenesis of chloroplast is the fact  
that some, if not most, of its components must come from the outside of the organelle itself. The  
import mechanisms must take into account to what part within the different sub-compartments  
the proteins are being targeted; hence the proteins being imported from the cytoplasm must be  
45 able to cross the different internal membrane barriers before they can reach their destinations.  
The import mechanism must also take into account how to tightly coordinate the interaction  
between the plastid and the nucleus such that both nuclear and plastidic components are  
expressed in a synchronous and orchestrated manner. Changes in the developmental and  
physiological conditions within or surrounding plant cells can consequently change this tight  
20 coordination and therefore change how import mechanisms are regulated as well. Manipulation  
of these conditions and modulation of expression of the import components and their function  
can have critical and global consequences to the development of the plant and to several  
biochemical pathways occurring outside the chloroplast. Expression patterns of such genes have  
been determined using microarray technology.

25 Microarray technology allows monitoring of gene expression levels for thousands of  
genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to  
glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US  
Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from  
such microarray experiments conducted with AFGC chips containing about 10,000 non-  
30 redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA  
of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in a mutant in a mutant (CiA2) of *Arabidopsis thaliana*, that is distributed in chloroplast biogenesis relative to wild type grown in the same conditions were identified, compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones that are involved in the import of proteins to chloroplast and chloroplast biogenesis.

Examples of genes and gene products that are involved in the import of proteins to chloroplast are shown in the Reference, Sequence, Protein Group, and Protein Group Matrix tables. While chloroplast protein import polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different chloroplast protein import responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Manipulation of one or more chloroplast protein import gene activities are useful to modulate the biological processes and/or phenotypes listed below. Chloroplast protein import responsive genes and gene products can act alone or in combination. Useful combinations include chloroplast protein import responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. Manipulation of one or more chloroplast protein import gene activities are useful to modulate the biological processes and/or phenotypes listed below.

Such chloroplast protein import responsive genes and gene products can function to either increase or dampen the above phenotypes or activities in response to changes in the regulation of import mechanisms. Further, promoters of chloroplast protein transport responsive

genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by chloroplast protein transport or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the chloroplast protein transport responsive genes when the desired sequence is operably linked to a promoter of a chloroplast protein transport responsive gene. The MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Chloroplast (relating to SMD 8093, SMD 8094)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Chloroplast genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Chloroplast Genes Identified By Cluster Analyses Of Differential Expression

##### Chloroplast Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Chloroplast genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Chloroplast (relating to SMD 8093, SMD 8094) of the MA\_diff and/or AFLP\_diff table(s).

##### Chloroplast Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Chloroplast genes. A group in the MA\_clust is considered a Chloroplast pathway or network if the group comprises a cDNA ID that also

appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

5

#### Chloroplast Genes Identified By Amino Acid Sequence Similarity

Chloroplast genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Chloroplast genes.

Groups of Chloroplast genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Chloroplast pathway or network is a group of proteins that also exhibits Chloroplast functions/utilities.

#### B.3.a. USE OF CHLOROPLAST PROTEIN IMPORT RESPONSIVE GENES TO MODULATE PHENOTYPES

Chloroplast protein import responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
  - Roots
  - Stems
  - Leaves
- Development
  - Plastid Biogenesis
  - Plastid Division
  - Plastid Development
  - Thylakoid Membrane Structures
- Differentiation
  - Plastid/Chloroplast Differentiation
- Photosynthesis
  - Carbon Dioxide Fixation

- Transport
  - Transcription/Translation Regulation Within Transport Complex
  - Phosphate Translocation
  - Targeted Starch Deposition And Accumulation
- Biosynthesis Of Essential Compounds
  - Lipid Biosynthesis
  - Riboflavin Biosynthesis
  - Carotenoid Biosynthesis
  - Aminoacid Biosynthesis

To improve any of the phenotype(s) above, activities of one or more of the chloroplast protein import responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Saito et al. (1994, Plant Physiol. 106: 887-95), Takahashi et al (1997, Proc. Natl. Acad. Sci. USA 94: 11102-07) and Koprivova et al. (2000, Plant Physiol. 122: 737-46).

#### B.3.b. USE OF CHLOROPLAST PROTEIN IMPORT-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the chloroplast protein import responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Growth and	• Regulation of Leaf	Reinbothe et al. (1997) Proc.

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Differentiation	Development Including Photosynthetic Apparatus	Natl. Acad. Sci. USA. 94:8890-8894 Eggink and Hooper (2000) J. Biol. Chem. 275:9087-9090 Jagtap et al. (1998) J Exptl Botany 49:1715-1721
	<ul style="list-style-type: none"> <li>Regulation of Plastid Biogenesis and Plastid Division</li> </ul>	Lawrence and Kindle (1997) J. Biol. Chem. 272:20357- 20363 Lahiri and Allison (2000) Plant Physiol. 123:883-894
	<ul style="list-style-type: none"> <li>Development of Plastid Inner/Outer and thylakoid Membrane Structures</li> </ul>	Kouranov et al. (1999) J. Biol. Chem. 274:25181- 25186 Jackson et al. (1998) J. Biol. Chem. 273:16583-16588 Li and Chen (1997) J. Biol. Chem. 272:10968-10974 Lawrence and Kindle (1997) J. Biol. Chem. 272:20357- 20363 Silva-Filho et al. (1997) J. Biol. Chem. 272:15264- 15269
	<ul style="list-style-type: none"> <li>Regulation of transcription and/or translation related to</li> </ul>	May and Soll (2000) Plant Cell 12:53-63 Caliebe et al. (1997) EMBO

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	maintenance of stability of protein-protein interaction within transport complex	J. 16:7342-7350
Physiology	<ul style="list-style-type: none"> <li>• Modulation of Photosynthesis</li> </ul>	Sung and Krieg (1979) Plant Physiol 64: 852-56
	<ul style="list-style-type: none"> <li>• Regulation of Lipid Biosynthesis</li> </ul>	Bourgis et al. (1999) Plant Physiol. 120:913-922 Reverdatto et al. (1999) Plant Physiol. 119:961-978 Roesler et al. (1997) Plant Physiol. 113:75-81
	<ul style="list-style-type: none"> <li>• Regulation of Riboflavin (Vitamin B) biosynthesis</li> </ul>	Jordan et al. (1999) J. Biol. Chem. 274:22114-22121
	<ul style="list-style-type: none"> <li>• Regulation of phosphate translocation across chloroplast membrane</li> </ul>	Flugge (1999) Annu. Rev. Plant Physiol. Plant Mol. Biol. 50:27-45  Silva-Filho et al. (1997) J. Biol. Chem. 272:15264-15269
	<ul style="list-style-type: none"> <li>• Regulation of targeted starch deposition and accumulation</li> </ul>	Yu et al. (1998) Plant Physiol. 116:1451-1460
	<ul style="list-style-type: none"> <li>• Modulation of protein targeting and translocation across chloroplast membrane</li> </ul>	Summer and Cline (1999) Plant Physiol. 119:575-584 Dabney-Smith et al. (1999) J. Biol. Chem. 274:32351-



GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		32359 Hinnah et al. (1997) EMBO J. 16:7351-7360
	<ul style="list-style-type: none"> <li>Regulation of carotenoid biosynthesis</li> </ul>	Bonk et al. (1996) Plant Physiol. 111:931-939
	<ul style="list-style-type: none"> <li>Regulation of amino acid biosynthesis</li> </ul>	Flugge (1999) Annu. Rev. Plant Physiol. Plant Mol. Biol. 50:27-45
	<ul style="list-style-type: none"> <li>Regulation of secondary metabolism</li> </ul>	Flugge (1999) Annu. Rev. Plant Physiol. Plant Mol. Biol. 50:27-45
Signal Transduction	<ul style="list-style-type: none"> <li>Regulation of gene transcriptional activity specific to chloroplast protein import</li> </ul>	Chen et al. (2000) Plant Physiol. 122:813-822. Macasev et al. (2000) Plant Physiol. 123:811-816
	<ul style="list-style-type: none"> <li>Regulation of protein target signal cleavage and protein degradation</li> </ul>	Lang et al. (1998) J. Biol. Chem. 273:30973-30978 Jackson et al. (1998) J. Biol. Chem. 273:16583-16588 Richter and Lamppa (1998) Proc. Natl. Acad. Sci. USA. 95:7463-7468
	<ul style="list-style-type: none"> <li>Regulation of ion channel conformation</li> </ul>	Van der Wijngaard and Vredenberg (1999) J. Biol.

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	and activity	Chem. 274:25201-25204
	<ul style="list-style-type: none"> <li>Regulation of kinase and phosphatases synthesis and activity</li> </ul>	<p>Waegemann and Soll (1996) J. Biol. Chem. 271:6545-6554</p> <p>Li et al. (2000) Science 287-300-303</p> <p>Muller et al. (2000) J. Biol. Chem. 275:19475-19481</p>
	<ul style="list-style-type: none"> <li>Modulation of Molecular Chaperone and Other Protein Folding Activity</li> </ul>	<p>Bonk et al. (1996) Plant Physiol. 111:931-939</p> <p>Walker et al. (1996) J. Biol. Chem. 271:4082-4085</p> <p>Kessler and Blobel (1996). Proc. Natl. Acad. Sci. USA 93:7684-7689</p> <p>Jackson et al. (1998) J. Biol. Chem. 273:16583-16588</p>

Other biological activities that can be modulated by the chloroplast protein import responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

- Chloroplast protein import responsive genes are characteristically differentially transcribed in response to fluctuating chloroplast protein import levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff reports the changes in transcript levels of various chloroplast protein import responsive genes that are differentially expressed among the mutants and the wild type.

Profiles of some of these chloroplast protein import responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	<p>Responders to defective chloroplast protein import</p> <p>Genes induced by defective import</p>	<ul style="list-style-type: none"> <li>Chloroplast protein import regulation</li> <li>Chloroplast protein import and transport</li> <li>Chloroplast import metabolism</li> <li>Synthesis of secondary metabolites and/or proteins</li> <li>Modulation of chloroplast import response transduction pathways</li> <li>Changes in chloroplast membranes</li> <li>Specific gene transcription initiation</li> <li>Chloroplast and</li> </ul>	<ul style="list-style-type: none"> <li>Transporters</li> <li>Metabolic enzymes</li> <li>Change in cell membrane structure and potential</li> <li>Kinases and phosphatases</li> <li>Transcription activators</li> <li>Change in chromatin structure and/or localized DNA topology</li> <li>Redox control</li> <li>Metabolic enzymes concerned with chloroplast biochemistry</li> <li>Organelle gene expression and translation</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		non-chloroplast metabolic pathways	
Down-regulated transcripts	<p>Responders to defective chloroplast protein import.</p> <p>Genes repressed by defective chloroplast protein import</p> <p>Genes with unsTable mRNAs when chloroplast import is defective</p> <p>Genes with discontinued expression or unsTable mRNA in presence of chloroplast protein import</p>	<ul style="list-style-type: none"> <li>Regulation of chloroplast protein import pathways released</li> <li>Chloroplast protein import and transport</li> <li>Chloroplast import metabolism</li> <li>Changes in pathways and processes operating in chloroplasts</li> <li>Changes in organelle membranes</li> <li>Loss of organelle gene expression, RNA and protein synthesis</li> <li>Changes in metabolism other</li> </ul>	<ul style="list-style-type: none"> <li>Transcription factors</li> <li>Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> <li>Change in chromatin structure and/or DNA topology</li> <li>Stability factors for protein mRNA synthesis and degradation</li> <li>Organelle transcription and translation proteins</li> <li>Metabolic enzymes</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		<p>than chloroplast protein import pathways</p> <ul style="list-style-type: none"> <li>• Chloroplast import metabolism</li> </ul>	

#### Use of Promoters of Chloroplast Genes

Promoters of Chloroplast genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Chloroplast genes where the desired sequence is operably linked to a promoter of a Chloroplast gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression

I.C. REPRODUCTION GENES, GENE COMPONENTS AND PRODUCTS

I.C.1. REPRODUCTION GENES, GENE COMPONENTS AND PRODUCTS

5 Reproduction genes are defined as genes or components of genes capable of modulating any aspect of sexual reproduction from flowering time and inflorescence development to fertilization and finally seed and fruit development. These genes are of great economic interest as well as biological importance. The fruit and vegetable industry grosses over \$1 billion USD a year. The seed market, valued at approximately \$15 billion USD annually, is even more lucrative.

10 Expression of many reproduction genes and gene products is orchestrated by internal programs or the surrounding environment of a plant, as described below. These genes and/or products have great importance in determining traits such as fruit and seed yield. Examples of such reproduction genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix tables, Knock-in, Knock-out, MA-diff and MA-clust. The  
15 biochemical functions of the protein products of many of these genes determined from comparisons with known proteins are also given in the Reference tables.

Reproduction Genes Identified by Phenotypic Observation

20 Reproduction genes were discovered and characterized from a much larger set of genes by experiments designed to find genes that cause phenotypic changes in flower, silique, and seed morphology. In these experiments, reproduction genes were identified by either (1) ectopic expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then cultivated and phenotypes, which varied from the parental "wild-type", were observed.

25 One particular example of reproductive genes are those that are regulated by AP2. AP2 is a transcription factor that regulates many genes, both as a repressor of some genes and as an activator of others. Some of these genes are those which establish the floral meristem or those which regulate floral organ identity and development. As such, AP2 has an effect on reproduction. This is, loss of AP2 activity is correlated with decreased male and female reproduction. AP2 is also known to have an effect on seed mass, and therefore on yield. That is,  
30 overexpression of AP2 is correlated with smaller seeds or seedless fruit while repression of AP2 correlates with larger seeds (see, e.g. US Patent No.: 5,994,622).

Another example of reproduction genes are those that are regulated by PISTILLATA (PI). PI is a transcription factor that regulates many genes both as a repressor and activator. Some of these genes are those which regulate floral organ identity and development, in conjunction with other transcription factors such as AP2 and AGAMOUS. As such, PI has an effect on reproduction in that loss of PI activity is correlated with decreased male reproduction. PI is also known to have an effect on carpel number, and therefore potentially on ovule/seed number and yield. Specifically, repression of PI results in increased carpel number and therefore ovule number.

Yet another example of reproductive genes are those that are regulated by MEDEA (MEA). MEA is a SET-domain containing protein that associates with other proteins to form complexes that affect chromatin structure and therefore gene expression. As such, loss of MEA function is correlated with global gene activation and repression leading to many phenotypes including decreased female reproduction and therefore reduced seed set and yield.

In the characterization of these and other reproduction genes, the following phenotypes were scored:

I. Flower

- Size
- Large
- Small
- Shape
- Abnormal organ numbers
- Agamous
- AP-2 like
- Color
- Number
- Fused Sepals

II. Silique

- Size
- Seed number
- Reduced
- Absent

- Seed color

The identified genes regulating reproduction are identified in the Knock-in and Knock-out Tables.

5 Reproduction Genes Identified by Differential Expression

Reproduction genes were also identified in experiments designed to discover genes whose mRNA products were in different concentrations in whole flowers, flower parts, and siliques relative to the plant as a whole. The MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108473, 108474, 108429, 108430, 108431, 108475, 108476, 108501). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Reproduction genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication

Reproduction Genes Identified By Cluster Analyses Of Differential Expression

Reproduction Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Reproduction genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108473, 108474, 108429, 108430, 108431, 108475, 108476, 108501 of the MA\_diff and/or AFLP\_diff table(s).

Reproduction Genes Identified By Correlation To Genes That Cause Physiological Consequences



Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Reproduction genes. A group in the MA\_clust is considered a Reproduction pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Reproduction Genes Identified By Amino Acid Sequence Similarity

Reproduction genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Reproduction genes. Groups of Reproduction genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Reproduction pathway or network is a group of proteins that also exhibits Reproduction functions/utilities.

It is assumed that the reproduction genes differentially expressed in floral parts and seeds are concerned with specifying flowers and seeds and their functions, and therefore modulations of such genes produce variant flowers and seeds.

Reproductive genes and gene products can function to either increase or dampen the phenotypes, biochemical activities and transcription profiles, either in response to changes of internal plant programs or to external environmental fluctuations.

#### C.1.a. USE OF REPRODUCTION GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE PHENOTYPES

The reproduction genes of the instant invention, when mutated or activated differently, are capable of modulating one or more processes of flower, seed and fruit development. They are thus useful for improving plants for agriculture and horticulture and for providing seeds with a better chemical composition for diverse industries including the food, feed and chemical industries. Reproduction genes, gene components and products are useful to alter the following traits and properties of plants:

##### I. Development

##### A. Flowering time and number of inflorescences

B. Flower development

- (a) Anther
- (b) Stamen
- (c) Pollen
- (d) Style
- (e) Stigma
- (f) Ovary
- (g) Ovule
- (h) Gametes

C. Pollination and Fertilization

- (a) Sporogenesis
- (b) Gametogenesis
- (c) Zygote formation
- (d) Embryo development
- (e) Endosperm development
- (f) Male sterility, hybrid breeding systems and heterosis

II. Cellular Properties:

- A. Cell size
- B. Cell shape
- C. Cell death
- D. Cell division
- E. Cell elongation
- F. Cell differentiation
- G. Meiosis

III. Organ Characteristics:

- A. Flowers
  - (1) Receptacle
  - (2) Sepals, Petals, and Tepals
    - a. Color

- b. Shape
- c. Size
- d. Number
- e. Petal Drop

(3) Androecium

Stamen

- (i) Anther
  - 1. Size
  - 2. Pollen
    - a. Sterile
    - b. Size, shape, weight, color
  - 3. Number
- (ii) Filament
  - 1. Size

(4) Gynoecium

- a. Carpel
    - (i) Ovary
    - (ii) Number
    - (iii) Length
  - b. Style
    - (i) Stigma
    - (ii) Ovule
- Size, shape, number

(5) Pedicel and Peduncle

- a. Size
- b. Shape

B. Seeds

- 1. Placenta
- 2. Embryo
- 3. Cotyledon
- 4. Endosperm

5. Suspensor
6. Seed coat (testa)
7. Aleurone
8. Development, including Apomixis (gametophytic,  
apospory, diplospory)
9. Dormancy
10. Germination

C. Fruits

1. Pericarp – thickness, texture
  - (a) Exocarp
  - (b) Mesocarp
  - (c) Endocarp
2. Development
  - (a) Seed set
  - (b) Fruit set
  - (c) False fruit
  - (d) Fruit elongation and maturation
  - (e) Dehiscence
3. Fruit drop

IV. Plant Seed Yield

1. Increased biomass
2. Better Harvest Index
3. Attraction of favorable insects
4. Better seed quality
5. Better yield of constituent chemicals

V. Plant Population Features:

A. Architecture

1. Shade avoidance
2. Planting density

To regulate any of the phenotype(s) above, activities of one or more of the reproduction genes or gene products can be modulated in an organism and tested by screening for the desired trait. Specifically, the gene, mRNA levels or protein levels can be altered in a plant using the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (*Methods Mol. Biol.* 82:259-266 (1998)) and/or screened for variants as in Winkler et al. (*Plant Physiol* 118:743-50 (1998)) and visually inspected for the desired phenotype or metabolically and/or functionally assayed.

#### C.1.b. USE OF REPRODUCTION GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the reproduction genes can be modulated to change biochemical or metabolic activities and/or pathways such as those examples noted below. Such biological activities can be measured according to the citations included in the Table below:

FUNCTION/PROCESS	EXAMPLES OF BIOCHEMICAL/MOLECULAR ACTIVITIES	Reference AND ASSAY
Metabolism	<p>Energy production and conversion</p> <ul style="list-style-type: none"> <li>- Glucosyl-transferase (CLONE_ID 1040)</li> <li>- Heme-binding protein (putative cytochrom B5) (CLONE_ID 3743)</li> </ul> <p>Storage protein synthesis</p> <p>Inorganic ion transport and metabolism</p> <ul style="list-style-type: none"> <li>- Peroxidase</li> </ul>	<p>Ap Rees, T. (1974). In "Plant Biochemistry. Biochemistry, Series One", Vol. 11. (H.L. Kornberg and D.C. Phillips, eds.), Butterworths, London.</p> <p>Juliano, B.O. and Varner, J.E. (1969). <i>Plant Physiol.</i> 44, 886-892.</p> <p>Bewley et al. (1993). <i>Plant Physiol. Biochem.</i> 31, 483-490.</p> <p>Hills, M. J. and Beevers, H. (1987). <i>Plant Physiol.</i> 84,</p>

	(CLONE_ID 100990)	272-276.
	- Cystathione beta synthase	Olsen, L. J. and Harada, J. J. (1991). In "Molecular Approaches to Compartmentalization and Metabolic Regulation (A. H. C. Huang and L. Taiz, eds.), ASPP, Rockville, Md.
	(CLONE_ID 21847)	
	Amino acid transport and metabolism	Mitsuhashi, W. and Oaks, A. (1994). Plant Physiol. 104, 401-407.
	- l-asparaginase	Walker-Smith, D. J., and Payne, J. W. (1985). Planta 164, 550-556.
	(CLONE_ID 92780)	
	- Putative peptide/amino acid transporter	Salmenkallio, M. and Sopanen, T. (1989). Plant Physiol. 89, 1285-1291.
	(CLONE_ID 113723)	Baumgartner, B. and Chrispeels, M. J. (1976). Plant Physiol. 58, 1-6.
	Carbohydrate transport and metabolism	Elpidina, E. N. et al. (1991). Planta 185, 46-52.
	- Glucose transport protein	Ericson, M. C. and Chrispeels, M. J. (1973). Plant Physiol. 52, 98-104.
	(CLONE_ID 33727)	Kern, R. and Chrispeels, M. J. 1978) Plant Physiol. 62, 815-819.
	- Putative sugar transporter	Dilworth, M. F. and Dure, L. III. (1978). Plant Physiol. 61, 698-702.
	(CLONE_ID 3250)	
	- Starch biosynthesis	
	Coenzyme metabolism	
	- Tyrosine aminotransferase	
	(ROOTY/SUPERROOT1)	
	(CLONE_ID 14570)	
	- Formate dehydrogenase	

	<p>(CLONE_ID 7530)</p> <p>Lipid metabolism</p> <ul style="list-style-type: none"> <li>- Branched chain <math>\alpha</math>-ketoacid dehydrogenase E2 subunit</li> </ul> <p>(CLONE_ID 25116)</p> <ul style="list-style-type: none"> <li>- Acyl carrier protein-1</li> </ul> <p>(CLONE_ID 14291)</p> <ul style="list-style-type: none"> <li>- Lipid metabolic enzymes</li> </ul> <p>Secretion</p> <ul style="list-style-type: none"> <li>- Sensor protein RcsC-like</li> </ul> <p>(CLONE_ID 16461)</p> <ul style="list-style-type: none"> <li>- Signal recognition particle</li> </ul> <p>RP54 (CLONE_ID 22158)</p>	<p>Chrispeels, M. J. and Jones, R. L. (1980/81). <i>Isr. J. Bot.</i> 29, 222-245.</p> <p>Gould, S. E. B., and Rees, D. A. (1964). <i>J. Sci. Food Agric.</i> 16, 702-709.</p>
Modulate floral organ number	<p>Transcriptional control</p> <p>ANT (AP2-domain) DNA binding protein</p> <p>SUP (Zinc finger)</p>	<p>Elliot et al. (1996). <i>Plant Cell</i> 8, 155-168.</p> <p>Sakai et al. (2000). <i>Plant Cell</i> 12, 1607-1618.</p> <p>Jacobsen and Meyerowitz (1997). <i>Science</i> 277, 1100-1103.</p>
Floral organ size	<p>Transcriptional control</p> <p>ANT (AP2-domain) DNA binding protein</p>	<p>Mizukami et al. (2000). <i>PNAS</i> 97, 942-947.</p> <p>Krizek (1999). <i>Developmental Genetics</i> 25,</p>

		224-236.
Female organ number/Floral meristem size	Membrane receptor kinase signal transduction CLV1 (LRR domain and kinase domain) receptor CLV2 (LRR domain) receptor  CLV3 (Receptor ligand)	Clark and Meyerowitz (1997). Cell 89, 575-585 Jeong et al. (1999). Plant Cell 11, 1925-1934. Fletcher et al. (1999). Science 283, 1911-1914.
Female reproduction	DNA binding protein AG (MADS domain) DNA binding protein	Yanofsky et al. (1990). Nature 346, 35-39.
Female reproduction	Signal transduction CTR1 (Raf kinase)	Kieber et al. (1993). Cell 72, 427-441.
Male organ number	DNA methylation MET1 (DNA methyltransferase)	Jacobsen and Meyerowitz (1997). Science 277, 1100- 1103.
Seed size control	DNA binding protein AP2 (AP2 domain) RAP2 (AP2 domain)	Jofuku et al. (1994). Plant Cell 6, 1211-1225. US Patent #6,093,874; #5,994,622
Seed size control	Polycomb group protein complex FIE, FIS2, MEA	Luo et al. (2000). PNAS 97, 10637-10642.



Seed size control	DNA methylation MET1	Scott et al. (2000). Development 127, 2493-2502. Vinkenoog et al. (2000). Plant Cell 12, 2271-2282. Luo et al. (2000). PNAS 97, 10637-10642.
Embryo development/Embryo viability	CAAT box binding complex LEC1/HAP3 HAP2, HAP5	Lotan et al. (1998). Cell 93, 1195. US Patent #6,235,975
Embryo development/Seed dormancy	DNA binding proteins ABI4 (AP2 domain)  FUS3 (B3 domain) VP1 (B3 domain)	Finkelstein et al. (1998). Plant Cell 10, 1043-1054. Luerssen et al. (1998). Plant J. 15, 755-764.
Embryo development	Signal transduction ABI1, ABI2 [Serine/threonine protein phosphatase 2C (PP2C)]	Leung et al. (1994). Science 264, 1448-1452. Leung et al. (1997). Plant Cell 9, 759-771.
Endosperm development	Chromatin level control of gene activity Polycomb complex; FIE, MEA, FIS2	Ohad et al. (1996). PNAS 93, 5319-5324. US Patent #6,229,064 Kiyosue et al. (1999). PNAS 96, 4186-4191.

		Grossniklaus et al. (1998). Science 280, 446-450. Chaudhury et al. (1997) PNAS 94, 4223-4228.
Integument development/Seed coat development	DNA binding AP2, ANT (AP2 domain)  BEL1 (Homeodomain)	Jofuku et al. (1994). Plant Cell 6, 1211-1225. Klucher et al. Plant Cell 8, 137-153. Reiser et al. (1995). Cell 83, 735-742.
Anthocyanin production	Secondary transporter TT12 (MATE; multidrug and toxic compound extrusion)	Debeaujon et al. (2001). Plant Cell 13, 853-872.
Anthocyanin production	DNA binding protein TT8 (Basic helix-loop-helix domain)	Nesi et al. (2000). Plant Cell 12, 1863-1878.
Fruit development	Chromatin level control of gene activity Polycomb complex; FIE, MEA, FIS2	Ohad et al. (1996). PNAS 93, 5319-5324. Kiyosue et al. (1999). PNAS 96, 4186-4191. Grossniklaus et al. (1998). Science 280, 446-450. Chaudhury et al. (1997) PNAS 94, 4223-4228.
Fruit size control	Signal transduction FW2.2 (c-Ras P21)	Frery et al. (2000). Science 289, 85-88.
Fruit development/Pod shattering	Transcriptional control SHP1, SHP2, FUL (MADS domain) DNA binding	Liljegren et al. (2000). Nature 404, 766-770.

	proteins	Ferrandiz et al. (2000). Science 289, 436-438..
Transcription and Posttranscription	<p>Transcription</p> <ul style="list-style-type: none"> <li>- SRF-domain AGL11 (CLONE_ID 32791)</li> <li>- AP2-domain containing protein (CLONE_ID 332)</li> <li>- Myb-DNA binding protein (CLONE_ID 94597)</li> </ul> <p>Transcription factors</p> <p>Signal transduction mechanisms</p> <ul style="list-style-type: none"> <li>- Protein-kinases</li> <li>- Phosphatases</li> <li>- meiosis proteins</li> <li>- Chromatin remodeling proteins</li> <li>- Chaperones</li> <li>- Chalcone synthase</li> <li>- Putative Ser/Thr protein kinase (CLONE_ID 31383)</li> <li>- ER6-like protein (implicated in ethylene signal transduction) (CLONE_ID 7474)</li> </ul>	<p>Delseny, M. et al. (1977). Planta 135, 125-128.</p> <p>Lalonde, L. and Bewley, J. D. (1986). J. Exp. Bot. 37, 754-764.</p> <p>Walling, L. et al. (1986). PNAS 83, 2123-2125.</p> <p>Okamuro, J. K. and Goldberg, R. B. (1989). In "Biochemistry of Plants, Vol 15." Academic Press, Inc.</p> <p>Wong, J. et al. (1995). Genes Dev. 9, 2696-2711.</p> <p>Dimitrov et al. (1994). J. Cell Biol. 126, 591-601.</p> <p>Landsberger, N. and Wolffe, A. P. (1997). EMBO J. 16, 4361-4373.</p> <p>Bogdanove, A. J. and Martin, G. G. (2000). PNAS 97, 8836-8840.</p> <p>Zhu, H. et al. Science July 26, 2001: 10.1126/science.1062191 (Reports).</p>

	<p>Translation, ribosomal structure and biogenesis</p> <ul style="list-style-type: none"> <li>- Ribosomal proTein S15A (CLONE_ID 17466)</li> <li>- Translation initiation factor (CLONE_ID 103464)</li> </ul> <p>Posttranslational modification, protein turnover, chaperones</p> <ul style="list-style-type: none"> <li>- DnaJ-domain containing protein (CLONE_ID 4150)</li> <li>- Cyclophilin-like protein (CLONE_ID 35643)</li> </ul>	
Cell division and Repair	<p>Cell division and chromosome partitioning</p> <ul style="list-style-type: none"> <li>- Protein of unknown function with tropomyosin-, myosin tail- and filament- domains (CLONE_ID 15546)</li> <li>- Actin-1 (CLONE_ID 25785)</li> </ul> <p>DNA replication,</p>	<p>Rogan, P. G. and Simon, E. W. (1975). New Phytol. 74, 273-275.</p> <p>Morahashi, Y. and Bewley, J. D. (1980). Plant Physiol 66, 70-73.</p> <p>Morahashi, Y. et al. (1981). Plant Physiol. 68, 318-323.</p> <p>Morahashi, Y. (1986). Physiol. Plant. 66, 653-658.</p> <p>Zlatanova, J. et al. (1987). Plant Mol. Biol. 10, 139-144.</p>

	<p>recombination and repair</p> <ul style="list-style-type: none"> <li>- Proliferating cell nuclear antigen-1 (axillary protein, DNA polymerase I delta)</li> <li>(CLONE_ID 28554)</li> <li>- AAA-type ATPase, cdc48</li> <li>(CLONE_ID 100292)</li> </ul> <p>Cell envelope biogenesis, outer membrane</p> <ul style="list-style-type: none"> <li>- dTDP-D-glucose 4,6-dehydratase</li> <li>(CLONE_ID 28597)</li> <li>- Putative cinnamoyl-CoA reductase</li> <li>(CLONE_ID 109228)</li> </ul>	<p>Zlatanova, J. and Ivanov, P. (1988). Plant Sci. 58, 71-76.</p>
--	--	---

Other biological activities that are modulated by the reproductive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table, for example.

C.1.c. USE OF REPRODUCTION GENES, GENE COMPONENTS AND PRODUCTS  
TO MODULATE TRANSCRIPTION LEVELS

Reproduction genes are characteristically differentially transcribed in response to cell signals such as fluctuating hormone levels or concentrations, whether internal or external to an organism or cell. Many reproduction genes belong to networks or cascades of genes under the control of regulatory genes. Thus some reproduction genes are useful to modulate the expression of other genes. Examples of transcription profiles of reproduction genes are described in the Table below with associated biological activities. "Up-regulated" profiles are those where the mRNA transcript levels are higher in flowers, flower parts or siliques as compared to the plant as a whole. "Down-regulated" profiles represent higher transcript levels in the whole plant as compared to flowers, flower parts or siliques alone.

TRANSCRIPT LEVELS	TYPE OF GENES WITH ALTERED ACTIVITY	PHYSIOLOGICAL CONSEQUENCES OF ALTERING GENE EXPRESSION	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENES WITH ALTERED EXPRESSION
Up Regulated Transcripts	Genes that control flower differentiation, number and size	Flowers form from flower meristem	Transcription Factors Signal transduction Membrane Structure Protein kinases Phosphatases Meiosis proteins Chromatin remodeling proteins Chaperones Chalcone synthase
Flower Reproduction Genes	Genes that promote petal, stamen and carpel formation	Floral organs mature	
	Genes controlling flower-specific metabolism such as petal pigments	Flavonoid pathways induced	
			Amino acid transport

TRANSCRIPT LEVELS	TYPE OF GENES WITH ALTERED ACTIVITY	PHYSIOLOGICAL CONSEQUENCES OF ALTERING GENE EXPRESSION	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENES WITH ALTERED EXPRESSION
	<p>Genes that promote ovule formation</p> <p>Genes that promote fertilization, seed, embryo and endosperm formation</p>		<p>and metabolism</p> <p>Storage protein synthesis</p> <p>Lipid metabolic enzymes</p> <p>Carbohydrate transport and metabolism</p> <p>Starch biosynthesis</p>
<p>AP2</p> <p>Reproduction Genes</p>	<p>Genes activated by AP2 transcription factors</p> <p>Genes that induce petal and stamen formation</p>	<p>Many steps and pathways induced, developmental and metabolic</p> <p>No petals or stamens produced</p>	<p>Proteins associated with:</p> <p>Energy production and conversion</p> <p>Amino acid transport and metabolism</p> <p>Carbohydrate transport and metabolism</p> <p>Lipid metabolism</p> <p>Transcription and</p>

TRANSCRIPT LEVELS	TYPE OF GENES WITH ALTERED ACTIVITY	PHYSIOLOGICAL CONSEQUENCES OF ALTERING GENE EXPRESSION	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENES WITH ALTERED EXPRESSION
			signal transduction Poor translational modification DNA replication Chromatin remodeling
Down-Regulated Transcripts Flower Reproduction Genes	Genes that repress flower development  Genes that induce stem, leaf and other organ differentiation  Genes that negatively regulate flower specific metabolism  Genes that negatively regulate ovule	Flowers form from flower meristem  Non-floral organs are repressed  Flower-specific pathways are induced	Transcription factors Signal transduction pathways Kinases and phosphatases  Chromatin remodeling proteins



TRANSCRIPT LEVELS	TYPE OF GENES WITH ALTERED ACTIVITY	PHYSIOLOGICAL CONSEQUENCES OF ALTERING GENE EXPRESSION	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENES WITH ALTERED EXPRESSION
AP2 Reproduction Genes	<p>formation, meiosis, fertilization and seed development</p> <p>Genes activated by AP2 transcription factors</p> <p>Genes that induce petal and stamen formation</p>	<p>Many steps and pathways induced, developmental and metabolic</p> <p>No petals or stamens produced</p>	<p>Proteins associated with:</p> <p>Energy production and conversion</p> <p>Amino acid transport and metabolism</p> <p>Carbohydrate transport and metabolism</p> <p>Lipid metabolism</p> <p>Transcription and signal transduction</p> <p>Poor translational modification</p> <p>DNA replication</p> <p>Chromatin remodeling</p>

While polynucleotides and gene products modulating reproduction can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations

include different polynucleotides and/or gene products of the instant invention that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of a polynucleotide and/or gene product(s) capable of modulating reproduction with a hormone responsive polynucleotide, particularly one  
5 affected by gibberellic acid and/or auxin, is also useful because of the interactions that exist between hormone regulated pathways, and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways.

#### Use Of Promoters And Reproduction Genes

Promoter of reproduction genes are useful for transcription of desired polynucleotides, both plant and non-plant. For example, extra copies of carbohydrate transporter genes can be operably linked to a reproduction gene promoter and inserted into a plant to increase the "sink"  
10 strength of flowers or siliques. Similarly, reproduction gene promoters can be used to drive transcription of metabolic enzymes capable of altering the oil, starch, protein or fiber of a flower or silique. Alternatively, reproduction gene promoters can direct expression of non-plant genes that can, for instance confer insect resistance specifically to a flower.  
15

I.C.2. OVULE GENES, GENE COMPONENTS AND PRODUCTS

The ovule is the primary female sexual reproductive organ of flowering plants. It contains the egg cell and, after fertilization occurs, contains the developing seed. Consequently, the ovule is at times comprised of haploid, diploid and triploid tissue. As such, ovule development requires the orchestrated transcription of numerous polynucleotides, some of which are ubiquitous, others that are ovule-specific and still others that are expressed only in the haploid, diploid or triploid cells of the ovule.

Although the morphology of the ovule is well known, little is known of these polynucleotides and polynucleotide products. Mutants allow identification of genes that participate in ovule development. As an example, the pistillata (PI) mutant replaces stamens with carpels, thereby increasing the number of ovules present in the flower. Accordingly, comparison of transcription levels between the wild-type and PI mutants allows identification of ovule-specific developmental polynucleotides.

Changes in the concentration of ovule-specific polynucleotides during development results in the modulation of many polynucleotides and polynucleotide products. Examples of such ovule-specific responsive polynucleotides and polynucleotide products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, AFLP\_diff, MA\_diff and/or AFLP\_diff, and MA\_clust tables. These polynucleotides and/or products are responsible for effects on traits such as fruit production and seed yield.

While ovule-specific developmentally responsive polynucleotides and polynucleotide products can act alone, combinations of these polynucleotides also affect fruit and seed growth and development. Useful combinations include different ovule-specific developmentally responsive polynucleotides and/or polynucleotide products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of an ovule-specific developmentally responsive polynucleotide and/or polynucleotide product with an environmentally responsive polynucleotide is also useful because of the interactions that exist between development, hormone regulated pathways, stress pathways and nutritional pathways. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in a common

pathway. The MA\_diff and/or AFLP\_diff, AFLP\_int and AFLP\_diff Table(s) report the transcript levels of the experiment (see EXPT ID: 108595). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Ovule genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Ovule Genes Identified By Cluster Analyses Of Differential Expression

Ovule Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Ovule genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108595 of the MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

#### Ovule Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Ovule genes. A group in the MA\_clust is considered a Ovule pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Ovule Genes Identified By Amino Acid Sequence Similarity

Ovule genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Ovule genes. Groups of Ovule

genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Ovule pathway or network is a group of proteins that also exhibits Ovule functions/utilities.

Such ovule-specific developmentally responsive polynucleotides and polynucleotide products can function to either increase or dampen the above phenotypes or activities either in response to transcript changes during ovule development or in the absence of ovule-specific polynucleotide fluctuations. More specifically, ovule-specific developmentally responsive polynucleotides and polynucleotide products are useful to or modulate one or more of the following phenotypes:

- Egg Cell
- Maturation
  - for development of parthenogenic embryos
- Metabolism
- Polar nuclei
- Fusion
  - for development of parthenogenic endosperm
- Central Cell
- Maturation
- Metabolism
  - For alteration of endosperm metabolism
- Synergids
- Maturation
- Programmed cell death
- Nucellus
- Maturation
- Integuments
- Maturation
- Funiculus
- Extension
  - for increased seed

- Cuticle
- Maturation
- Tensile properties
  - for increased seed size
- Ovule
- Modulation of ovule senescence
- Shaping
  - for increased seed number

To produce the desired phenotype(s) above, one or more of the ovule-specific developmentally responsive polynucleotides and polynucleotide products can be tested by screening for the desired trait. Specifically, the polynucleotide, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Weigel et al. (2000, Plant Physiol 122: 1003-14) and Winkler et al. (1998, Plant Physiol 118: 743-50).

Alternatively, the activities of one or more of the ovule-specific developmentally responsive polynucleotides and polynucleotide products can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
Cell Growth and Differentiation	-Programmed Cell Death	Pennell and Lamb (1997) Plant Cell 9, 1157-1168
	- DNA Methylation and	Adams et al. (2000)

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
	Imprinting	Development 127: 2493-502
Organ Growth and Development	<ul style="list-style-type: none"> <li>-Ovule Growth and Development</li> <li>-Ethylene Response</li> <li>-Megagametophyte Development</li> <li>-Seed Growth and Development</li> <li>-Fertilization Independent Seed Development</li> </ul>	<p>De Martinis and Mariani (1999) Plant Cell 11: 1061-72 Christensen et al. (1997) Sexual Plant Reproduc 10: 49-64</p> <p>Scott et al. (1998) Development 125: 3329-41 Ohad et al. (1996) PNAS USA 93: 5319-24 Chaudhury et al. (1997) PNAS USA 94: 4223-28</p>
Signal Transduction	<ul style="list-style-type: none"> <li>-Ethylene Metabolism</li> <li>- Protein Remodeling</li> <li>-Sucrose Mobilization and Partitioning</li> <li>-Pollen Tube Adhesion</li> <li>-Jasmonic Acid</li> </ul>	<p>DeMartinis and Mariani (1999) Plant Cell 11: 1061-1072 Winkler et al. (1998) Plant Physiol 118: 743- 750</p>

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
	Biosynthesis -	
Senescence and Cell Death	-Apomixis	
Environmental Responses	-Wound and Defense Response Gene Expression  -Stress Response	Epple and Bohlmann (1997) Plant Cell 9: 509- 20 He et al. (1998) Plant J. 14: 55-63

Other biological activities that can be modulated by the ovule-specific developmentally responsive polynucleotides and polynucleotide products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table section.

Ovule-specific developmentally responsive polynucleotides are characteristically differentially transcribed in response to fluctuating developmental-specific polynucleotide levels or concentrations, whether internal or external to a cell. The MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) report the changes in transcript levels of various ovule-specific developmentally responsive polynucleotides in ovules.

These data can be used to identify a number of types of ovule-specific developmentally responsive polynucleotides. Profiles of these different ovule-specific developmentally responsive polynucleotides are shown in the Table below with examples of associated biological activities.



TRANSCRIPTS AFFECTED BY	TYPES OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Ethylene Signals	Responders to Ethylene	- Ethylene Perception -Ethylene Uptake - Modulation of Ethylene Response Transduction Pathways - Specific Gene Transcription Initiation	-Transcription Factors -Transporters
Protein Remodeling		-Repression of Pathways to Optimize Abscissic acid Response Pathways	-Inhibit Transport of Abscissic acid -Degradation
Lower at 1 hours than 6 hours	High Abscissic acid Responders		-Abscissic acid Metabolic Pathways
	Repressor of Abscissic acid Deprivation Pathways	Negative Regulation of Abscissic acid Pathways	

#### Use of Promoters of Ovule Genes

Promoters of Ovule genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Ovule genes where the desired sequence is operably linked to a promoter of a Ovule gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce

antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

I.C.3. SEED AND FRUIT DEVELOPMENT GENES, GENE COMPONENTS  
AND PRODUCTS

The ovule is the primary female sexual reproductive organ of flowering plants. At maturity it contains the egg cell and one large central cell containing two polar nuclei encased by two integuments that, after fertilization, develops into the embryo, endosperm, and seed coat of the mature seed, respectively. As the ovule develops into the seed, the ovary matures into the fruit or silique. As such, seed and fruit development requires the orchestrated transcription of numerous polynucleotides, some of which are ubiquitous, others that are embryo-specific and still others that are expressed only in the endosperm, seed coat, or fruit. Such genes are termed fruit development responsive genes.

Changes in the concentration of fruit-development responsive polynucleotides during development results in the modulation of many polynucleotides and polynucleotide products. Examples of such fruit development responsive polynucleotides and polynucleotide products relative to leaves and floral stem are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, AFLP\_diff, MA\_diff and/or AFLP\_diff, MA\_clust, Knock-in and Knock-out tables. The polynucleotides were discovered by isolating fruits at developmental stages from Arabidopsis wild-type ecotype "Wassilewskija", and measuring the mRNAs expressed in them relative to those in a leaf and floral stem sample. These polynucleotides and/or products are responsible for effects on traits such as seed size, seed yield, seed composition, seed dormancy, fruit ripening, fruit production, and pod shattering.

While fruit development responsive polynucleotides and polynucleotide products can act alone, combinations of these polynucleotides also affect fruit and seed growth and development. Useful combinations include different polynucleotides and/or polynucleotide products that have similar transcription profiles or similar biological activities, and members of the same or functionally similar biochemical pathways. In particular, modulation of transcription factors and/or signal transduction pathways are likely to be useful for manipulating whole pathways and hence phenotypes. In addition, the combination of ovule-developmentally responsive polynucleotides and/or polynucleotide products with environmentally responsive polynucleotides is also useful because of the interactions that exist between development, hormone regulated pathways, stress and pathogen induced pathways and nutritional pathways. Here, useful

combinations include polynucleotides that may have different transcription profiles, and participate in common or overlapping pathways but combine to produce a specific, phenotypic change.

Such fruit development responsive polynucleotides and polynucleotide products can function to either increase or dampen the above phenotypes or activities either in response to transcript changes in fruit development or in the absence of fruit development polynucleotide fluctuations.

The MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) report the transcript levels of the experiment (see EXPT ID: 108436, 108437, 108438). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Fruit genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the AFLP\_diff and/or MA\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Fruit Genes Identified By Cluster Analyses Of Differential Expression

#### Fruit Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Fruit genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108436, 108437, 108438 of the MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

#### Fruit Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Fruit genes. A group in the MA\_clust is considered

a Fruit pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Fruit Genes Identified By Amino Acid Sequence Similarity

Fruit genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Fruit genes. Groups of Fruit genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Fruit pathway or network is a group of proteins that also exhibits Fruit functions/utilities.

- Use of fruit development responsive genes to modulate phenotypes

Manipulation of the polynucleotides in the mature ovule, developing embryo, endosperm, seed coat and fruit enables many features of seed and fruit to be improved including the following:

- Female fertility, megasporogenesis, embryo and endosperm development, ovule size, endosperm size, embryo size, seed size, seed yield, seed protein, seed oil, seed starch, seed cell number, cell size, seed coat development, organ size, dormancy and acquisition of desiccation tolerance, seed storage and longevity, seed germination, apomixis, production of seedless fruit and vegetables and hybrid seed production.

To improve any of the phenotype(s) above, activities of one or more of the fruit development responsive polynucleotides and polynucleotide products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the polynucleotide, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed.

**Use of fruit development responsive genes to modulate biochemical activities**

The activities of one or more of the fruit-expressed polynucleotides and polynucleotide products can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological changes can be achieved and measured according to citations such as the following:

1. Winkler et al. (1998). Plant Physiol. 118, 743-750
2. Weigel et al. (2000). Plant Physiol. 122, 1003-1014
3. Cosgrove (1997). Plant Cell 9, 1031-1041
4. Jacobs (1997). Plant Cell 9, 1021-1029
5. Reismeier et al. (1994). EMBO J. 13, 1-7
6. Carland et al. (1999). Plant Cell 11, 2123-2138
7. Cheng et al. (1996). Plant Cell 8, 971-983
8. Weber et al. (1995). Plant Cell 7, 1835-1846
9. Leyser and Furer (1992). Development 116, 397-403
10. Hayashi et al. (1998). Plant Cell 10, 183-196.
11. Pyke (1999). Plant Cell 11, 549-556
12. Lotan et al. (1998). Cell 93, 1195-1205
13. Lending and Larkins (1989). Plant Cell 1, 1011-1023
14. Hong et al. (1996). Development 122, 2051-2058.
15. Fernandez et al. (2000). Science 289, 436-438
16. D'Aoust et al. (1999). Plant Cell 11, 2407-2418
17. Bewley (1997). Plant Cell 9, 1055-1066
18. Heath et al. (1986). Planta 169, 304-312
19. Browse et al. (1986). Anal. Biochem. 152, 141-145
20. D'Aoust et al. (1999). Plant Cell 11, 2407-2418

Other biological activities that can be modulated by the fruit-specific developmentally responsive polynucleotides and polynucleotide products are listed in Reference Tables. Assays for detecting such biological activities are described in the table as well as in the Protein Domain tables.

	BIOLOGICAL FUNCTION	POTENTIAL UTILITY	CITATION	ASSAY	CITATION
<b>Ovule Growth, Ovule</b>	Ethylene and ethylene signal	Manipulate female fertility.	De Martinis and Mariani (1999). Plant	Analyze ovule and seed	Winkler et al. (1998). Plant Physiol. 118,

<b>Development and Seed Growth and Development</b>	transduction pathway	Manipulate megasporogenesis.	Cell 11, 1061-1072. Silencing gene expression of the ethylene-forming enzyme results in a reversible inhibition of ovule development in transgenic tobacco plants.	development by light microscopy or by confocal microscopy.	743-750. Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis.
	Examples: AP2 domain DNA binding proteins; EREBP, EBF	Manipulate female gametophyte development.	enzyme results in a reversible inhibition of ovule development in transgenic tobacco plants.	Test for fertilization independent endosperm development.	Weigel et al. (2000). Plant Physiol 122, 1003-1014.
	Example: Leucine-rich receptor kinase; ETR-like	Manipulate fertilization independent endosperm development.	Christensen et al. (1997). Sexual Plant Reproduc. 10, 49-64.	Test for fertilization independent embryo development.	Activation tagging in Arabidopsis. Ohad et al. (1996). PNAS USA 93, 5319-5324. A mutation that allows endosperm development without fertilization
	Example: Raf kinase; CTR	Manipulate embryo development.	Megagametogenesis in Arabidopsis wild type and the Gf mutant.	Test for fertilization independent seed production.	Chaudhury et al. (1997). PNAS USA 94, 4223-4228. Fertilization-
		Manipulate fertilization independent seed development.	Christiansen and Drews, unpublished	Analyze seed size.	
		Manipulate ovule size.		Analyze seed yield.	
		Manipulate endosperm size.		Analyze seed composition.	
		Manipulate		Analyze fruit size.	

		embryo size.  Manipulate seed size.  Manipulate seed yield.  Manipulate seed protein.  Manipulate seed oil.  Manipulate starch production.  Manipulate cell number.  Manipulate cell size.  Produce seedless fruit and vegetables  Manipulate fruit size.			independent seed development in Arabidopsis thaliana  De Martinis and Mariani (1999). Plant Cell 11, 1061-1072. Silencing gene expression of the ethylene-forming enzyme results in a reversible inhibition of ovule development in transgenic tobacco plants. Christensen et al. (1997). Sexual Plant Reproduc. 10, 49-64. Megagametogenesis in Arabidopsis
--	--	---	--	--	--



					<p>wild type and the Gf mutant.</p> <p>Scott et al. (1998). Development 125, 3329-3341. Parent-of-origin effects on seed development in <i>Arabidopsis thaliana</i></p> <p>Heath et al. (1986). <i>Planta</i> 169, 304-312.</p> <p>Browse et al. (1986). <i>Anal. Biochem.</i> 152, 141-145.</p> <p>D'Aoust et al. (1999). <i>Plant Cell</i> 11, 2407-2418.</p>
	<b>2. Growth and developmental control</b>	Manipulate female fertility.	Wilson et al. (1996). <i>Plant Cell</i> 8, 659-671. A	Analyze ovule and seed development by light	Winkler et al. (1998). <i>Plant Physiol.</i> 118, 743-750.

	<p><b>genes</b></p> <p>----</p> <p><i>Upregulated genes</i></p> <p>Example: DNA binding proteins; tiny-like, AGL1, FBP2, AGL9, AP3, CPC-like myb.</p> <p>Example: Protein kinase; ASK1.</p> <p>Example: Auxin conjugating enzyme; indole-3-acetate beta-glucosyltransferase.</p> <p>Example: S/T protein kinase; APK1.</p> <p>Example:</p>	<p>Manipulate megasporogenesis.</p> <p>Manipulate female gametophyte development.</p> <p>Manipulate fertilization independent endosperm development.</p> <p>Manipulate fertilization independent embryo development.</p> <p>Manipulate fertilization independent seed development.</p> <p>Manipulate ovule size.</p> <p>Manipulate endosperm size.</p>	<p>dissociation insertion causes a semidominant mutation that increases expression of TINY, an Arabidopsis gene related to APETALA2.</p> <p>Zhao et al (1999).</p> <p>Developmental Genetics 25, 209-223. The ASK1 gene regulates development and interacts with the UFO gene to control floral organ identity in Arabidopsis.</p> <p>Flanagan et al. (1996).</p> <p>Plant J. 10, 343-53.</p>	<p>microscopy or by confocal microscopy.</p> <p>Test for fertilization independent endosperm development.</p> <p>Test for fertilization independent embryo development.</p> <p>Test for fertilization independent seed production.</p> <p>Analyze seed size.</p> <p>Analyze seed yield.</p> <p>Analyze seed composition.</p> <p>Analyze fruit size.</p> <p>Analyze seedling size.</p> <p>Analyze</p>	<p>Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis.</p> <p>Weigel et al. (2000). Plant Physiol 122, 1003-1014.</p> <p>Activation tagging in Arabidopsis.</p> <p>Ohad et al. (1996). PNAS USA 93, 5319-5324. A mutation that allows endosperm development without fertilization</p> <p>Chaudhury et al. (1997). PNAS USA 94, 4223-4228.</p> <p>Fertilization-</p>
--	--	--	---	--	--

	<p>Leucine-rich receptor kinase; CLV1, ER, BRI, Cf-2-like.</p> <p>-----</p> <p><i>Downregulate d genes</i></p> <p>Example: Cyclin-dependent kinase; cdc2.</p>	<p>Manipulate embryo size.</p> <p>Manipulate organ size and number.</p> <p>Manipulate seed size.</p> <p>Manipulate seed yield.</p> <p>Manipulate seedling size through seed size.</p> <p>Manipulate seedling vigor through seed size.</p> <p>Manipulate seed protein.</p> <p>Manipulate seed oil.</p> <p>Manipulate starch production.</p> <p>Manipulate integument development.</p> <p>Manipulate</p>	<p>Specific expression of the AGL1 MADS-box gene suggests regulatory functions in Arabidopsis gynoecium and ovule development.</p> <p>Angenent et al. (1994). Plant J 1994. 5, 33-44. Co-suppression of the petunia homeotic gene fbp2 affects the identity of the generative meristem.</p> <p>AGL9 web page.</p> <p>Wada et al. (1997) Science 277, 1113-6.</p> <p>Epidermal cell</p>	<p>seedling viability.</p> <p>Screen for changes in shatter time.</p> <p>Screen for changes in germination frequency.</p> <p>Screen for seed longevity and viability.</p>	<p>independent seed development in Arabidopsis thaliana</p> <p>De Martinis and Mariani (1999). Plant Cell 11, 1061-1072. Silencing gene expression of the ethylene-forming enzyme results in a reversible inhibition of ovule development in transgenic tobacco plants.</p> <p>Christensen et al. (1997). Sexual Plant Reproduc. 10, 49-64.</p> <p>Megagametoge</p>
--	---	--	--	---	---

		seedcoat development.	differentiation in Arabidopsis determined by a Myb homolog		nesis in Arabidopsis wild type and the Gf mutant.
		Manipulate cell size.	homolog		Scott et al.
		Manipulate cell number.	CPC.		(1998).
		Manipulate homeotic gene expression.	Szerszen et al. (1994).		Development 125, 3329-3341. Parent-of-origin effects on seed development in Arabidopsis thaliana.
		Manipulate organ size.	Science 16, 1699-1701.		Heath et al. (1986).
		Manipulate meristem size.	iaglu, a gene from Zea mays involved in conjugation of growth hormone indole-3-acetic acid.		Planta 169, 304-312.
		Produce seedless fruit and vegetables	Ito et al. (1997).		Browse et al. (1986).
		Manipulate fruit size.	Plant Cell Physiol. 38, 248-258.		Biochem. 152, 141-145.
		Manipulate time of seed dispersal.	A serine/threonine protein kinase gene isolated by an in vivo binding		D'Aoust et al. (1999).
		Manipulate seed viability upon storage.			Plant Cell 11, 2407-2418.
		Manipulate germination			

		frequency.	procedure using the Arabidopsis floral homeotic gene product, AGAMOUS.  Clark et al. (1997). Cell 89, 575-585.  The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis.  Torii et al. (1996). Plant Cell 8, 735- 746. The Arabidopsis ERECTA gene encodes a putative receptor		
--	--	------------	---	--	--

			protein kinase with extracellular leucine-rich repeats.  Li and Chory (1997). Cell 90, 929-38. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction.		
	<b>3. Cell senescence and cell death</b>  Example: Cystatin  Example: WIPK	Manipulate female fertility.  Manipulate seed set.  Manipulate seed yield.  Manipulate seed size.  Manipulate fruit set.  Promote	Solomon et al. (1999). Plant Cell 11, 431-444. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in	Analyze ovule and seed development by light microscopy or by confocal microscopy.  Analyze seed set.  Analyze seed size.  Analyze seed	Winkler et al. (1998). Plant Physiol. 118, 743-750. Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis. Weigel et al. (2000). Plant

		apomixis.  Produce seedless fruit and vegetables.	plants.  Zhang et al. (2000). Plant J. 23, 339-347. Multiple levels of tobacco WIPK activation during the induction of cell death by fungal elicitors.	yield.  Analyze fruit set. Screen for fertiliza-tion independent seed development.	Physiol 122, 1003-1014. Activation tagging in Arabidopsis. Ohad et al. (1996). PNAS USA 93, 5319-5324. A mutation that allows endosperm development without fertilization
	<b>4. Protein remodelin g</b>  Example: DNA-J protein/chaperones	Manipulate female fertility.  Manipulate female gametophyte development.  Promote apomixis.  Manipulate endosperm development.  Manipulate	Christensen et al. (1997). Sexual Plant Reproduc. 10, 49-64. Megagametogenesis in Arabidopsis wild type and the Gf mutant. Cory Christiansen and Gary Drews,	Test for altered female fertility, seed set, seed yield.  Analyze ovule development by light microscopy or by confocal microscopy.  Analyze seed size.	Winkler et al. (1998). Plant Physiol. 118, 743-750. Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis. Weigel et al. (2000). Plant Physiol 122,

		embryo development.  Manipulate seed size.  Manipulate seed yield.  Manipulate seed protein.  Manipulate seed oil.  Manipulate starch.  Produce seedless fruit and vegetables.	unpublished	Analyze seed yield.  Analyze seed composition.	1003-1014.  Activation tagging in Arabidopsis.  Christensen et al. (1997). Sexual Plant Reproduc. 10, 49-64.  Megagametogenesis in Arabidopsis wild type and the Gf mutant.  Ohad et al. (1996). PNAS USA 93, 5319-5324. A mutation that allows endosperm development without fertilization  Scott et al. (1998). Development 125, 3329-3341. Parent-of-origin
--	--	--	-------------	--	--



					effects on seed development in Arabidopsis thaliana.  Heath et al. (1986). Planta 169, 304-312.  Browse et al. (1986). Anal. Biochem. 152, 141-145.  D'Aoust et al. (1999). Plant Cell 11, 2407-2418.
	<p><i>5. Sucrose mobilization and partitioning</i></p> <p>Example: Invertase inhibitor</p> <p>Example: bZIP transcription factor (translation of</p>	<p>Manipulate female fertility.</p> <p>Manipulate ovule development.</p> <p>Manipulate seed development.</p> <p>Manipulate endosperm</p>	<p>Mapping of tomato genes associated with sugar metabolism.</p> <p>Tomato Genetics Co-op Report 48, 22-23 (1998)</p> <p>Ikeda et al. (1999). Plant Physiol 121,</p>	<p>Analyze ovule and seed development by light microscopy or by confocal microscopy.</p> <p>Determine female fertility.</p> <p>Analyze seed</p>	<p>Winkler et al. (1998). Plant Physiol. 118, 743-750.</p> <p>Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis.</p> <p>Weigel et al. (2000). Plant</p>

	<p>bZIP protein is inhibited by sucrose levels greater than 25 mM)</p> <p>Example: Lipoxxygenase</p> <p>----</p> <p><i>Downregulate d gene</i></p> <p>Example: SNF1-related protein kinase</p>	<p>development.</p> <p>Manipulate embryo development.</p> <p>Manipulate seed size.</p> <p>Manipulate seed yield.</p> <p>Manipulate seed protein.</p> <p>Manipulate seed oil.</p> <p>Manipulate starch.</p> <p>Manipulate cell size.</p> <p>Manipulate cell number.</p> <p>Manipulate organ size.</p> <p>Manipulate meristem size.</p> <p>Manipulate seedling size through seed size.</p> <p>Manipulate</p>	<p>813-820.</p> <p>Sucrose and Cytokinin</p> <p>Modulation of WPK4, a Gene</p> <p>Encoding a SNF1-Related Protein</p> <p>Kinase from Wheat.</p> <p>Rook et al. (1998). Plant J. 15, 253-263. Sucrose-specific signaling represses translation of the Arabidopsis ATB2 bZIP transcription factor gene.</p> <p>Rook et al. (1998). Plant Mol Biol 37,171-178.</p> <p>The light-regulated</p>	<p>mass.</p> <p>Analyze seed yield.</p> <p>Analyze seed composition.</p> <p>Analyze organ size.</p> <p>Analyze seedling size.</p> <p>Analyze seedling viability.</p>	<p>Physiol 122, 1003-1014.</p> <p>Activation tagging in Arabidopsis.</p> <p>Christensen et al. (1997). Sexual Plant Reproduc. 10, 49-64.</p> <p>Megagametogenesis in Arabidopsis wild type and the Gf mutant.</p> <p>Ohad et al. (1996). PNAS USA 93, 5319-5324. A mutation that allows endosperm development without fertilization</p> <p>Scott et al. (1998). Development 125, 3329-</p>
--	--	--	--	--	--

		seedling viability through seed size. Produce seedless fruit and vegetables. Translational control of gene expression in ovule and seed by sucrose. Manipulate assimilate partitioning in ovule and seed development.	Arabidopsis bZIP transcription factor gene ATB2 encodes a protein with an unusually long leucine zipper domain. Bunker et al. (1995). Plant Cell 7, 1319-1331. Sink limitation induces the expression of multiple soybean vegetative lipooxygenase mRNAs while the endogenous jasmonic acid level remains low. Lowry et al. (1998). Plant	3341. Parent-of-origin effects on seed development in Arabidopsis thaliana. 6. Heath et al. (1986). Planta 169, 304-312. 7. Browse et al. (1986). Anal. Biochem. 152, 141-145. 8. D'Aoust et al. (1999). Plant Cell 11, 2407-2418.
--	--	--	---	---

			Physiol. 116, 923-933. Specific soybean lipoxygenases localize to discrete subcellular compartments and their mRNAs are differentially regulated by source-sink status.		
	<p>6. <b><i>Jasmonic acid biosynthesis and signal transduction pathway</i></b></p> <p>Example: Biosynthetic enzyme; FMN oxidoreductase 12-</p>	<p>Targeted death of cells belonging to the female gametophyte, ovule or integuments. Delay senescence of unfertilized female gametophyte, ovule or integuments.</p>	<p>Sanders et al. (2000). Plant Cell 12, 1041-1062. The Arabidopsis DELAYED DEHISCENCE1 gene encodes an enzyme in the jasmonic acid synthesis pathway. Vijayan et al.</p>	<p>Test for altered female fertility. Analyze male fertility. Screen for enhanced expression of pathogen defense response genes.</p>	<p>Winkler et al. (1998). Plant Physiol. 118, 743-750. Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis Weigel et al. (2000). Plant Physiol 122,</p>

	oxophyto- dienoate reductase, OPR1, OPR1- like. Example: Signal transduction pathway kinase WIPK.	Manipulate female fertility Coordinate female with male reproduction. Manipulate male fertility. Enhanced defense response in ovules and seed	(1998). A role for jasmonate in pathogen defense of Arabidopsis. PNAS USA 95, 7209- 7214. Seo et al. (1999). Plant Cell 11, 289- 298. Jasmonate- based wound signal transduction requires activation of WIPK, a tobacco mitogen- activated protein kinase.		1003-1014. Activation tagging in Arabidopsis.
<i>Environment al responses</i>	<i>1. Wound and defense response gene</i>	Pathogen resistant ovules. Pathogen	Song et al. (1995). Science 270, 1804-1806. A	Resistance to Xanthomonas sp. Resistance to	Winkler et al. (1998). Plant Physiol. 118, 743-750.

	<p><b>expression</b></p> <p>Example: Leucine rich receptor S/T kinase; Xa21- like and TMK-like.</p> <p>Example: Cell wall- associated protein kinase WAK1.</p> <p>Example: Thionins.</p>	<p>resistant seeds.</p> <p>Pathogen resistant fruit.</p>	<p>receptor kinase-like protein encoded by the rice disease resistance gene, Xa21.</p> <p>Seo et al. (1995). Science 270, 1988- 1992. Tobacco MAP kinase: a possible mediator in wound signal transduction pathways.</p> <p>He et al. (1998). Plant J. 14, 55-63.</p> <p>Requirement for the induced expression of a cell wall associated receptor kinase for survival during the pathogen response.</p>	<p>known arabidopsis pathogens in ovules, seed and fruit.</p>	<p>Systematic reverse genetics of transfer-DNA- tagged lines of Arabidopsis.</p> <p>Weigel et al. (2000). Plant Physiol 122, 1003-1014.</p> <p>Activation tagging in Arabidopsis.</p> <p>Eppl and Bohlmann (1997). Plant Cell 9, 509- 520.</p> <p>Overexpress- ion of an endogenous thionin enhances resistance of Arabidopsis against Fusarium oxysporum.</p> <p>He et al.</p>
--	--	--	--	---	---

			<p>He et al. (1999). Plant Mol. Biol. 39, 1189-1196. A cluster of five cell wall- associated receptor kinase genes, Wak1-5, are expressed in specific organs of Arabidopsis.</p> <p>Eppl and Bohlmann (1997). Plant Cell 9, 509- 520.</p> <p>Overexpressi on of an endogenous thionin enhances resistance of Arabidopsis against Fusarium oxysporum.</p> <p>Ichimura et al.</p>	<p>(1998). Plant J. 14, 55-63.</p> <p>Requirement for the induced expression of a cell wall associated receptor kinase for survival during the pathogen response.</p>
--	--	--	---	---

			(1998). DNA Res. 5,341-5348. Molecular cloning and characterization of three cDNAs encoding putative mitogen-activated protein kinase kinases (MAPKKs) in Arabidopsis thaliana.		
	<p><b>2. Stress response to cold, drought, salinity, seed maturation, embryo development, ABA.</b></p> <p>Example: Dehydrins</p> <p>Example: NPK1-like protein kinase</p>	<p>Manipulate drought resistance.</p> <p>Manipulate desiccation tolerance in flowers, ovules and seeds.</p> <p>Manipulate cold tolerance in flowers, ovules, and seeds.</p>	<p>Close, T.J. (1996). Physiol.Plant 97, 795-803.</p> <p>Dehydrins: emergence of a biochemical role of a family of plant dehydration proteins.</p> <p>Kovtun et al. (2000). PNAS</p>	<p>Test for enhanced sensitivity to drought, dessication, cold, salinity, in ovules, developing seed and seedlings.</p> <p>Test for enhanced tolerance to drought,</p>	<p>Winkler et al. (1998). Plant Physiol. 118, 743-750.</p> <p>Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis.</p> <p>Weigel et al. (2000). Plant Physiol 122,</p>



197

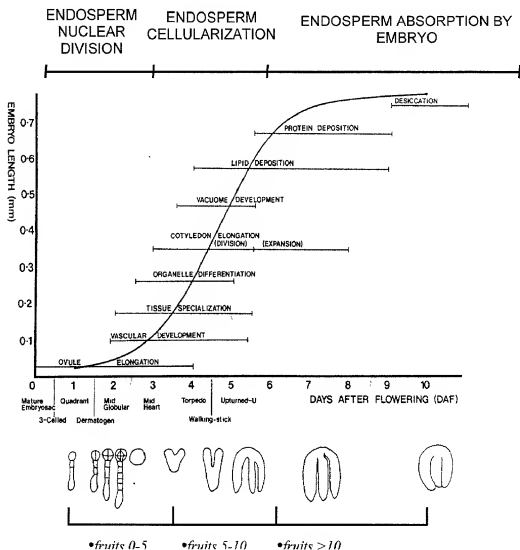
	acetate beta-glucosyltransferase.				
<b>Cell metabolism</b>	<p><b><i>Stearoyl-acyl carrier protein desaturase</i></b></p> <p>Example: C18 fatty acid desaturation</p>	<p>Production of oils high in saturated fatty acids</p> <p>Manipulate membrane composition</p>	<p>Merlo et al. (1998). Plant Cell 10, 1603-1621.</p>	<p>Analyze seed size.</p> <p>Analyze seed yield.</p> <p>Analyze seed composition.</p> <p>Analyze seed oil by gas chromatography.</p>	<p>Winkler et al. (1998). Plant Physiol. 118, 743-750.</p> <p>Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis.</p> <p>Weigel et al. (2000). Plant Physiol 122, 1003-1014.</p> <p>Activation tagging in Arabidopsis.</p> <p>Browse et al. (1986). Anal. Biochem. 152, 141-145.</p>
	<p><b>2. Manipulate</b></p>	<p>Manipulate asparagine</p>	<p>Mathews and Van Holde</p>	<p>Analyze seed size.</p>	<p>Winkler et al. (1998). Plant</p>

	<p><b><i>nitrogen economy</i></b></p> <p>Example: Asparaginase</p>	<p>degradation in ovules and seeds.</p> <p>Manipulate endosperm production.</p> <p>Manipulate embryo development.</p> <p>Manipulate ovule size.</p> <p>Manipulate seed size.</p>		<p>Analyze seed yield.</p> <p>Analyze seed composition.</p>	<p>Physiol. 118, 743-750.</p> <p>Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis.</p> <p>Weigel et al. (2000). Plant Physiol 122, 1003-1014.</p> <p>Activation tagging in Arabidopsis.</p> <p>Heath et al. (1986). Planta 169, 304-312.</p> <p>Browse et al. (1986). Anal. Biochem. 152, 141-145.</p> <p>D'Aoust et al. (1999). Plant Cell 11, 2407-2418.</p>
--	--	--	--	---	---

Fruit development responsive polynucleotides are characteristically differentially

transcribed in response to fluctuating developmental-specific polynucleotide levels or other signals, whether internal or external to a cell. MA\_diff and/or AFLP\_diff and/or AFLP\_diff reports the changes in transcript levels of various fruit development responsive polynucleotides in fruits.

These data can be used to identify a number of types of fruit development responsive polynucleotides. Profiles of some of these different fruit development responsive polynucleotides are shown in the table below with examples of the kinds of associated biological activities. Because development is a continuous process and many cell types are being examined together, the expression profiles of genes overlap between stages of development in the chart below.



Transcript Levels	Developmental Process	Metabolic Pathways	Examples of Biochemical Activity
<p>(0-5 mm)&gt;&gt;(5-10 mm) <math>\equiv</math> (&gt;10 mm)</p> <p>(0-5 mm)&gt;&gt;(5-10 mm) &gt; (&gt;10 mm)</p> <p>(0-5 mm)&gt;(5-10 mm) <math>\equiv</math> (&gt;10 mm)</p>	<p>Ovule Elongation</p> <p>Tissue Specialization</p> <ul style="list-style-type: none"> <li>- Vascular system</li> <li>- Meristem</li> <li>- Endosperm</li> <li>- Seed coat</li> <li>- Fruit</li> </ul>	<ul style="list-style-type: none"> <li>- Hormone Production, Transport, Perception, Signalling, Response (e.g., Gibberellin, Ethylene, Auxin)</li> <li>- Cell wall Biosynthesis</li> <li>- Lipid Biosynthesis</li> <li>- Specific Gene Transcription Initiation</li> <li>- Sucrose Mobilization and Partitioning</li> <li>- Sucrose Signaling</li> <li>- Lipoxigenase Localization</li> <li>- Repressors of Metabolic Pathways</li> <li>- Protein Remodeling</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Changes in cytoskeletal protein activity modulating cell structure</li> <li>- Stability factors for protein translation</li> <li>- Changes in cell wall/membrane structure</li> <li>- Chromatin structure and/or DNA topology</li> <li>- Biosynthetic enzymes</li> </ul>
<p>(5-10 mm) &gt;&gt;(0-5 mm) &gt; (&gt;10 mm)</p> <p>(5-10 mm) &gt;(0-5 mm) <math>\equiv</math> (&gt;10 mm)</p> <p>(5-10 mm) &gt;&gt;(0-5 mm) <math>\equiv</math> (&gt;10 mm)</p>	<p>Tissue Specialization</p> <ul style="list-style-type: none"> <li>- Vascular System</li> </ul> <p>Organelle Differentiation</p> <p>Cotyledon Elongation (cell division)</p> <p>Vacuome Development</p> <p>Lipid Deposition</p>	<ul style="list-style-type: none"> <li>- Cell Wall Biosynthesis</li> <li>- Specific Gene Transcription Initiation</li> <li>- Sucrose Mobilization and Partitioning</li> <li>- Sucrose Signaling</li> <li>- Repressors of Metabolic Pathways</li> <li>- Auxin Perception, Response and Signaling</li> <li>- Protein Remodeling</li> <li>- Lipid Biosynthesis and Storage</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> <li>- Changes in cytoskeletal protein activity modulating cell structure</li> <li>- Stability of factors for protein translation</li> <li>- Changes in cell wall/membrane structure</li> <li>- Chromatin structure and/or DNA topology</li> <li>- Biosynthetic enzymes</li> </ul>

Transcript Levels	Developmental Process	Metabolic Pathways	Examples of Biochemical Activity
(>10 mm) >(0-5 mm) $\equiv$ (5-10 mm)	Cotyledon Elongation (expansion) Lipid Deposition Protein Deposition Desiccation	<ul style="list-style-type: none"> <li>- Cell Elongation</li> <li>- Specific Gene Transcription Initiation</li> <li>- Sucrose Mobilization and Partitioning</li> <li>- Sucrose Signaling</li> <li>- Lipoygenase Localization</li> <li>- Repressors of metabolic pathways</li> <li>- Hormone Perception, Response and Signaling (e.g. abscisic acid)</li> <li>- Protein Remodeling</li> <li>- Protein synthesis and Storage</li> <li>- Lipid Synthesis and Storage</li> <li>- Acquisition of Dessication Tolerance</li> <li>- Senescence</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones for protein translation</li> <li>- Changes in cell wall/membrane structure</li> <li>- Chromatin structure and/or DNA topology</li> <li>- Biosynthetic enzymes</li> <li>- Metabolic enzymes</li> </ul>
(0-5 mm) < (5-10 mm) $\equiv$ (>10 mm)  (0-5 mm) << (5-10 mm) $\equiv$ (>10 mm)  (0-5 mm) << (5-10 mm) <(>10 mm)  (0-5 mm) << (5-10 mm) < (5-10 mm)	Ovule Elongation -Repressors of Ethylene production Tissue specialization - Vascular System - Meristem - Cotyledon - Seed Coat	<ul style="list-style-type: none"> <li>- Cell elongation</li> <li>- Negative regulation of ethylene pathways</li> <li>- Maintenance of Ethylene response</li> <li>- Changes in pathways and processes operation in cells</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> <li>- Stability of factors</li> <li>- Biosynthetic enzymes</li> <li>- Metabolic enzymes</li> </ul>
(5-10 mm) < (0-5 mm) $\equiv$ (>10 mm)	Organelle differentiation Cotyledon elongation (division) Vacuome development Lipid development Desiccation	<ul style="list-style-type: none"> <li>- Negative regulation of hormone pathways</li> <li>- Maintenance of hormone response</li> <li>- Changes in pathways and processes operation in cells</li> <li>- Dehydration and acquisition of desiccation tolerance</li> <li>- Senescence</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> </ul>
(>10 mm) <(0-5	Cotyledon Elongation	- Cell elongation	- Transcription

Transcript Levels	Developmental Process	Metabolic Pathways	Examples of Biochemical Activity
mm) $\equiv$ (5-10 mm)	(expansion) Lipid deposition Protein deposition Desiccation	<ul style="list-style-type: none"> <li>- Negative regulation of hormone pathways</li> <li>- Maintenance of hormone response</li> <li>- Changes in pathways and processes operation in cells</li> <li>- Dehydration and acquisition of desiccation tolerance</li> <li>- Senescence</li> </ul>	<ul style="list-style-type: none"> <li>Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> <li>- Metabolic enzymes</li> <li>- Biosynthetic enzymes</li> </ul>
(0-5 mm) $\equiv$ (5-10 mm) $\equiv$ (>10 mm)	All stages	<ul style="list-style-type: none"> <li>- Ribosome/polysome production and maintenance</li> <li>- Housekeeping genes</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> </ul>

I.D. DEVELOPMENT GENES, GENE COMPONENTS AND PRODUCTS

I.D.1. IMBIBITION AND GERMINATION RESPONSIVE GENES, GENE  
COMPONENTS AND PRODUCTS

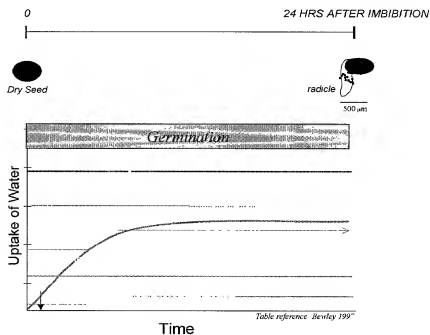
Seeds are a vital component of the world's diet. Cereal grains alone, which comprise ~90% of all cultivated seeds, contribute up to half of the global per capita energy intake. The primary organ system for seed production in flowering plants is the ovule. At maturity, the ovule consists of a haploid female gametophyte or embryo sac surrounded by several layers of maternal tissue including the nucleus and the integuments. The embryo sac typically contains seven cells including the egg cell, two synergids, a large central cell containing two polar nuclei, and three antipodal cells. that . Pollination results in the fertilization of both egg and central cell. The fertilized egg develops into the embryo. The fertilized central cell develops into the endosperm. And the integuments mature into the seed coat. As the ovule develops into the seed, the ovary matures into the fruit or silique. Late in development, the developing seed ends a period of extensive biosynthetic and cellular activity and begins to desiccate to complete its development and enter a dormant, metabolically quiescent state. Seed dormancy is generally an undesirable characteristic in agricultural crops, where rapid germination and growth are required. However, some degree of dormancy is advantageous, at least during seed development. This is particularly true for cereal crops because it prevents germination of grains while still on the ear of the parent plant (preharvest sprouting), a phenomenon that results in major losses to the agricultural industry. Extensive domestication and breeding of crop species have ostensibly reduced the level of dormancy mechanisms present in the seeds of their wild ancestors, although under some adverse environmental conditions, dormancy may reappear. By contrast, weed seeds frequently mature with inherent dormancy mechanisms that allow some seeds to persist in the soil for many years before completing germination.

Germination commences with imbibition, the uptake of water by the dry seed, and the activation of the quiescent embryo and endosperm. The result is a burst of intense metabolic activity. At the cellular level, the genome is transformed from an inactive state to one of intense transcriptional activity. Stored lipids, carbohydrates and proteins are catabolized fueling seedling



growth and development. DNA and organelles are repaired, replicated and begin functioning. Cell expansion and cell division are triggered. The shoot and root apical meristem are activated and begin growth and organogenesis. Schematic 4 summarizes some of the metabolic and cellular processes that occur during imbibition. Germination is complete when a part of the embryo, the radicle, extends to penetrate the structures that surround it. In Arabidopsis, seed germination takes place within twenty-four (24) hours after imbibition. As such, germination requires the rapid and orchestrated transcription of numerous polynucleotides. Germination is followed by expansion of the hypocotyl and opening of the cotyledons. Meristem development continues to promote root growth and shoot growth, which is followed by early leaf formation.

*Schematic 4. Cellular and Molecular Events in Seed Imbibition and Germination*



Genes with activities relevant to imbibition-germination and early seedling growth are described in the two sections A and B below.

#### D.2.a. IDENTIFICATION OF IMBIBITION AND GERMINATION GENES

Imbibition and germination includes those events that commence with the uptake of water by the quiescent dry seed and terminate with the expansion and elongation of the shoots and roots. The germination period exists from imbibition to when part of the embryo, usually the radicle, extends to penetrate the seed coat that surrounds it.

Imbibition and germination genes identified herein are defined as genes, gene components and products capable of modulating one or more processes of imbibition and germination described above. They are useful to modulate many plant traits from early vigor to yield to stress tolerance. Examples of such germination genes and gene products are shown in the Reference and Sequence Tables. The functions of many of the genes were deduced from comparisons with known proteins and are also given in the REF Tables.

##### Imbibition and Germination Genes Identified by Phenotypic Observations

Imbibition and germination genes are active, potentially active or more active during growth and development of a dry seed into a seedling. These genes herein were discovered and characterized from a much larger set of genes in experiments designed to find genes that cause poor germination.

In these experiments, imbibition and germination genes were identified by either 1) ectopic expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The seeds were then imbibed and cultivated under standardized conditions and any phenotypic differences in the modified plants compared with the parental "wild-type" seedlings were recorded. The genes causing the changes were deduced from the cDNA inserted or gene mutated. The phenotypic differences observed were poor germination and aberrant seedlings.

##### Imbibition and Germination Genes Identified by Differential Expression

Germination genes were also identified by measuring the relative levels of mRNA products of genes in different stages of germination of a seed versus the plant as a whole. Specifically, mRNA was isolated from whole imbibed seeds of Arabidopsis plants 1, 2, 3 or 4

days after imbibition and compared to mRNA isolated from dry seed-utilizing microarray procedures. The MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) report the transcript levels of the experiment. For transcript levels that were higher in the imbibed seed than in dry seed a "+" is shown. A "-" is shown when the transcript levels in dry seed were greater than those in imbibed seed. For more experimental detail, see the examples below:

Germination associated genes can be identified by comparing expression profiles of imbibed gibberellin treated and untreated ga1 mutant seed. Germination associated genes can also be identified by comparing expression profiles in late maturation seed from wild-type and mutants that are defective for the establishment of dormancy and can germinate precociously (e.g. aba1, aba2, abi4 in arabidopsis and vp1, vp5 in maize) or are defective for the specification of cotyledon identity and desiccation tolerance (e.g. lec1, lec2, and fus3).

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108461, 108462, 108463, 108464, 108528, 108529, 108530, 108531, 108545, 108546, 108547, 108518, 108529, 108543, 108544). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Imbibed & Germinating Seeds genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Imbibed & Germinating Seeds Genes Identified By Cluster Analyses Of Differential Expression

#### Imbibed & Germinating Seeds Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Imbibed & Germinating Seeds genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108461, 108462, 108463, 108464, 108528, 108529, 108530, 108531, 108545, 108546, 108547, 108518, 108529, 108543, 108544 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

#### Imbibed & Germinating Seeds Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Imbibed & Germinating Seeds genes. A group in the MA\_clust is considered a Imbibed & Germinating Seeds pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Imbibed & Germinating Seeds Genes Identified By Amino Acid Sequence Similarity

Imbibed & Germinating Seeds genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Imbibed & Germinating Seeds genes. Groups of Imbibed & Germinating Seeds genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Imbibed & Germinating Seeds pathway or network is a group of proteins that also exhibits Imbibed & Germinating Seeds functions/utilities.

D.2.b. USE OF IMBIBITION AND GERMINATION GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE PHENOTYPES

Imbibition and germination genes and gene products can be divided into those that act during primary events, secondary events, and/or termination. The genes and gene products of the instant invention are useful to modulate any one or more of the phenotypes described below:

I. Primary events

A. Dormancy

Imbibition and germination genes and gene products of the invention can act to modulate different types of dormancy including:

1. Primary dormancy – dormancy is established during seed development
2. Seed coat-imposed dormancy – dormancy is imposed by blocking water uptake, mechanical restraint of embryo, blocking the exit of inhibitors
3. Embryo dormancy – cotyledon mediated inhibition of embryonic axis growth
4. Secondary dormancy – dormancy is induced when dispersed, mature seeds are exposed to unfavorable conditions for germination (e.g. anoxia, unsuitable temperature or illumination).
5. Hormone-induced

B. Dormancy-breaking signal perception and transduction

Germination genes and gene products include those that are able to modulate the response to dormancy releasing signals such as:

1. Fruit ripening and seed development
2. Imbibition
3. Temperature – low and high (range 0-23°)
4. Light – particularly for coat imposed dormancy
  - a. White light
  - b. Intermittent illumination
  - c. Orange and red region of the spectrum (longer than 700 or 730 nm)

- d. Phytochrome
- 5. Coat softening
- 6. Chemicals
  - a. Respiratory inhibitors
  - b. Sulfhydryl compounds
  - c. Oxidants
  - d. Nitrogenous compounds
  - e. Growth regulators – GA, BA, ethylene
  - f. Various, ethanol, methylene blue, ethyl ether, fusicoccin
- 7. Oxygen and Carbon dioxide
- 8. Stress

## II. Secondary Events

During the secondary events of germination, dormancy-maintaining metabolism is repressed, dormancy-breaking metabolism is induced and structures surrounding the embryo weaken (where present). Germination genes and gene products are useful to modulate processes of the secondary events as follows:

- A. Water uptake
  - 1. Cell expansion
  - 2. Change in osmotic state
    - 1. Ion exchange
- B. Respiration – oxygen consumption

The genes and genes products of the invention can regulate the following pathways which resume during the first respiratory burst of germination:

- 1. Glycolysis
  - 2. Pentose phosphate
  - 3. Citric acid
  - 4. Tricarboxylic acid cycle
- C. Mitochondrial development

Tissues of the mature dry seed contain mitochondria, and although these organelles are poorly differentiated as a consequence of the drying, they contain sufficient Krebs cycle enzymes and terminal oxidases to provide adequate amount of ATP to support metabolism for several hours after imbibition. During germination of embryos, there appears to be two distinct patterns of mitochondrial development. In starch-storing seeds, repair and activation of preexisting organelles predominate, whereas oil-storing seeds typically produce new mitochondria. Germination genes and gene products of the invention are useful to modulate the repair, activation and biogenesis pathways of mitochondria. Specific examples are as follows:

1. Membrane formation and repair
2. DNA repair and synthesis
3. Protein synthesis
4. Coordinated regulation of mitochondrial and nuclear genomes

#### D. Metabolism

In addition to respiration and organelle activity, enzyme activity, DNA repair, RNA synthesis and protein synthesis are fundamental cellular activities intimately involved in the completion of germination and the preparation for subsequent growth. Imbibition and germination genes and gene products of the invention can participate in or modulate these activities. Examples are:

1. ABA response (for more detail see section on ABA-responsive genes)
2. GA response (for more detail see section on GA-responsive genes)
3. ATP synthesis and Adenylate Energy Charge during germination
4. The synthesis and utilization of reducing power: pyridine nucleotides (NADH and NADPH)

### III. Termination

The last stage of seed germination is characterized by the emergence of the radicle or root apex through the seed coat. Typically, the cell walls loosen and the radicle extends from the embryo during late germination. Germination genes and gene products are useful to modulate

the mobilization of stored reserves, DNA synthesis and cell division that are typical of this stage of germination.

To regulate any of the phenotype(s) above, activities of one or more of the late germination genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, *Methods. Mol. Biol.* 82:259-266) and/or screened for variants as in Winkler et al. (1998) *Plant Physiol* 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Dolan et al. (1993, *Development* 119: 71-84), Dolan et al. (1997, *Development* 124: 1789-98), Crawford and Glass (1998, *Trends Plant Science* 3: 389-95), Wang et al. (1998, *PNAS USA* 95: 15134-39), Gaxiola et al. (1998, *PNAS USA* 95: 4046-50), Apse et al. (1999, *Science* 285: 1256-58), Fisher and Long (1992, *Nature* 357: 655-60), Schneider et al. (1998, *Genes Devel* 12: 2013-21) and Hirsch (1999, *Curr Opin Plant Biol.* 2: 320-326).

#### D.2.c. USE OF IMBIBITION AND GERMINATION GENES, GENE COMPONENTS AND PRODUCT TO MODULATE BIOCHEMICAL ACTIVITIES

The roles of the biochemical changes associated with imbibition and germination can be appreciated from a summary of the processes occurring.

#### Physiology

Water plays an important role throughout the plant life cycle. The most dramatic example of this is in seed germination. Although germination is triggered by water, the germination response is also positively regulated by the plant growth regulators the gibberellins and negatively affected by the growth regulator abscisic acid. Genes that are activated by water and genes that are activated by gibberellins can be identified through expression profiling experiments using arabidopsis mutants defective for gibberellin biosynthesis or perception (*gal*, *gai*), abscisic acid biosynthesis or perception (*aba1*, *abi3*, and *abi4*) in the presence or absence of exogenous gibberellins. These genes can be used to promote seedling growth and development and other phases of plant development.



### Transcriptional Control of Gene Activity

At the end of seed development, dessication and dormancy have imposed a global state of repression on gene activity throughout the seed. Reactivation of the genome requires water and gibberellins. One function of the genes that are activated early by imbibition is the rapid and dramatic reversal of gene repression. For example, expression-profiling experiments revealed that several thousand genes are hyperactivated in arabidopsis upon imbibition. These include genes involved in metabolic pathways, genes that promote cell growth and division, and transcriptional control genes. Thus one class of genes expressed early in imbibition includes those that promote high levels of gene expression. Other early genes are responsible for regulating specific metabolic, cell, and developmental processes. The strategy for distinguishing these functions was outlined in the Introduction.

### Mobilization of Storage Reserves

In contrast to the synthesis and accumulation of reserves during seed development an important function of genes expressed during imbibition and germination is the control of the mobilization and catabolism of seed storage reserves in the endosperm (in grasses and cereals) and the embryo. The mobilization of seed storage reserves is triggered by imbibition and may occur over several days. There are three classes of high molecular weight seed storage reserves: carbohydrates, triacylglycerols, and storage proteins. Upon imbibition seed storage reserves are converted into forms that can be transported and metabolized. Genes encoding enzymes for storage reserve catabolism are expressed shortly after imbibition. Starch for example is converted to sucrose. Triacylglycerols are converted into acetyl-CoA. Storage proteins are converted into amino acids or deaminated to provide carbon skeletons for oxidation.

### Carbohydrate Catabolism

Starch is the most common storage carbohydrate in seeds. The primary components of starch are amylose and amylopectin.

### Mobilization

There are two pathways for starch catabolism – hydrolytic and phosphorolytic. The product of these pathways is the monosaccharide glucose. Examples of the enzymes responsible for hydrolytic catabolism of starch are: amylase, glucosidase, amylase, dextrinase, isoamylase. The enzyme responsible for phosphorolytic activity is starch phosphorylase.

#### Transport

The mobilization of starch involves the synthesis of sucrose from glucose, which can then be transported to sites for growth in the root and shoot. In some seeds, maltose may be a major form of transported carbohydrate. The production of sucrose-6-P from glucose involves the following enzymes: UDP-glucose pyrophosphorylase, sucrose-6-P synthetase, and sucrose phosphatase.

#### Sucrose Catabolism

In target tissues sucrose is hydrolyzed by fructofuransidase (invertase) and/or sucrose synthetase. The synthesis of glucose from glucose-1-P involves sucrose synthetase.

#### Cell Biology

The lumen of the endoplasmic reticulum (ER) is target for other hydrolase activities including mannosidase, glucosaminidase, acid phosphatase, phosphodiesterase, and phospholipase D.

### TRIACYLGLYCEROL (TAG) CATABOLISM

Triacylglycerols are the major storage lipids of seeds. The products of TAG catabolism in imbibed and germinating seed are glycerol and free fatty acids. Most of the glycerol is converted to sucrose for export. Free fatty acids are catabolized through oxidation through the glyoxylate cycle and gluconeogenesis.

#### Mobilization

Hydrolysis of triacylglycerols is by lipases yielding glycerol and free fatty acids. Free fatty acids are oxidized to acetyl-CoA and propionyl-CoA via oxidation requiring ATP and coenzyme A. Catabolism of unsaturated fatty acids also requires cis, trans-isomerases,

epimerases, and hydratases. Acetyl-CoA is oxidized through the citric acid cycle to CO<sub>2</sub> and H<sub>2</sub>O. More importantly, acetyl-CoA can be utilized via the glyoxylate cycle and gluconeogenesis for glucose synthesis. Free fatty acids are also broken down via oxidation. Glycerol is converted via phosphorylation and oxidation to DHAP and G3P, which are used to synthesize glucose or oxidized via the citric acid cycle. Examples of other induced enzymes include isocitrate lyase and malate synthetase

#### Transport

Most of the glycerol, acetyl-CoA, and propionyl-CoA are converted to sucrose for transport. This requires the enzymes glycerol kinase and glycerol phosphate oxidoreductase.

#### Cell Biology

Glyoxysome biogenesis is required to support fatty acid catabolism and gluconeogenesis. Upon exposure to light there is a loss of glyoxysomes due to their conversion to peroxisomes.

### STORAGE PROTEIN CATABOLISM

#### Mobilization

The hydrolysis of storage proteins to amino acids is performed by a diverse group of proteinases and peptidases. The peptidases include endopeptidases, aminopeptidases, and carboxypeptidases. They include the A and B class proteinases. The liberated amino acids are available for protein synthesis, for deamination and reutilization of ammonia via glutamine and asparagine synthesis, and to provide carbon skeletons for respiration. Several enzymes including, deaminase, asparagine synthetase, glutamine synthetase and glutamate dehydrogenase are important players in the mobilization and utilization of stored nitrogen in imbibed seed.

#### Transport

The major transported form of amino acid in germinated seeds is asparagine. In some species glutamine and/or homoserine are the major form of transported amino acid. Aspartate, glutamate, alanine, glycine, and serine can be converted to sucrose and transported as sucrose. Other amino acids are transported unchanged.

### Cell Biology

Proteinases are sequestered in lumen of endoplasmic reticulum (ER) which then fuses with protein bodies.

While catabolism is high in the storage tissues of imbibed seed the products of catabolism are transported to sites of growth including the shoot and root apices fueling respiration, biosynthesis, cell division and differentiation.

### DEVELOPMENT

Imbibition triggers several key processes for seedling development. One is the activation of the shoot and root apical meristems. The shoot apical meristem is responsible for two primary growth activities. One is the production of the protoderm, procambium and ground meristem. The protoderm gives rise to the epidermal system of the plant, the procambium to the primary vascular tissues, and the ground meristem to the ground tissues including the cortex and pith. The second is the production of leaf primordia, which arise on the flanks of the apex. Thus, activation of the shoot apical meristem results in shoot growth and organogenesis.

The root apical meristem, by contrast is responsible for vegetative root development. The first primary growth activity of the root apical meristem is the production of the protoderm, procambium and ground meristem. The second primary growth activity is the production of the cells that give rise to the root cap.

Genes that govern shoot apical meristem activation and development can be identified in arabidopsis by gene profiling experiments comparing gene expression in wild-type imbibed seed and partial loss-of-function *stm* (shootmeristemless) mutants (see SAM). Genes governing root meristem activity can be identified by gene profiling experiments comparing gene expression in wild-type imbibed seed and *rml* (rootmeristemless) mutants.

Genes identified in this way are useful to promote or retard meristem growth, modify and strengthen shoot and root development, promote leaf development as described below.

Changes in the concentration of imbibition-germination activated polynucleotides result in the modulation of many other polynucleotides and polynucleotide products. Examples of such activated responsive polynucleotides and polynucleotide products relative to leaves and floral stem and to fruits at different development stages are shown in the Reference and Sequence Tables. These polynucleotides and/or products are responsible for effects on traits such as

seedling growth, seedling viability, and seedling vigor. The polynucleotides were discovered by isolating seeds from Arabidopsis wild-type ecotype "Wassilewskija" imbibed for 24 hours, and measuring the mRNAs expressed in them relative to those in a leaf and floral stem sample and to those in fruits at different developmental stages.

While imbibition-germination activated polynucleotides and polynucleotide products can act alone, combinations of these polynucleotides also affect germination. Useful combinations include different polynucleotides and/or polynucleotide products that have similar transcription profiles or similar biological activities, and members of the same or functionally similar biochemical pathways. In addition, the combination of imbibition germination activated polynucleotides and/or polynucleotide products with environmentally responsive polynucleotides is also useful because of the interactions that exist between development, hormone-regulated pathways, stress and pathogen induced pathways and nutritional pathways. Here, useful combinations include polynucleotides that may have different transcription profiles, and participate in common or overlapping pathways but combine to produce a specific, phenotypic change.

Such imbibition and germination activated polynucleotides and polynucleotide products can function to either increase or dampen the above phenotypes or activities either in response to transcript changes in fruit development or in the absence of fruit-specific polynucleotide fluctuations.

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS ALTERED	CITATIONS INCLUDING ASSAYS
Growth , Differentiation and Development	Farnesylation Mediated Seed Dormancy	Pei et al (1998) Science 282: 287-290; Cutler et al. (1996) Science 273: 1239
Metabolic activity	Nitrogen metabolism	Goupil et al (1998) J Exptl Botany 49:1855-62
Metabolic activity	-H <sup>+</sup> export and membrane hyperpolarization	Cerana et al. (1983)
Metabolic activity	Chloroplast functioning	Benkova et al (1999) Plant

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS ALTERED	CITATIONS INCLUDING ASSAYS
		Physil 121: 245-252
Growth, Differentiation and development	Regulation of Morphogenesis	Riou-Khamlichi et al. (1999) Science 283: 1541-44
Metabolic activity	Cell Death	Lohman et al. (1994) Physiol Plant 92: 322-328
Growth and development	Promotion of cell division Shoot formation in absence of exogenous cytokinin	Kakimoto (1996) Science 274: 982-985
Metabolic activity	Membrane repair	Heath et al. (1986) Planta 169: 304-12 Browse et al. (1986) Anal Biochem 152: 141-5 D'Aoust et al (1999) Plant Cell 11: 2407-18
Metabolism	Organic molecule export	Moody et al. (1988) Phytochemistry 27: 2857-61
Metabolic activity	Nutrient Uptake	Uozumio et al. (2000) Plant Physiol 122: 1249-59
Metabolic activity	Ion export	Uozumi et al. (2000) Plant Physiol 122: 1249-59 Frachisse et al. (2000) Plant J 21: 361-71
Growth, Differentiation and development	Division and/or elongation	Zhang and Forde (1998) Science 279: 407-409. Coruzzi et al. U.S. Pat. No. 5,955,651

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS ALTERED	CITATIONS INCLUDING ASSAYS
Metabolic activity	Regulation of Molecular chaperones	Wisniewski et al. (1999) Physiolgia Plantarum 105: 600-608
Metabolic activity	Reactivation of Aggregation and Protein Folding	Lee and Vierling (2000) Plant Physiol. 122: 189-197
Metabolic activity	Maintenance of Native Conformation (cytosolic proteins)	Queitsch et al. (2000) The Plant Cell 12: 479-92
Metabolic activity	Regulation of Translational Efficiency	Wells et al. (1998) Genes and Development 12: 3236-51
Metabolic activity	DNA Repair	Bewley (1997) Plant Cell 9: 1055-66
Metabolic activity	Protein Synthesis using stored or newly synthesized mRNAs	Heath et al. (1986) Planta 169: 304-12
Metabolic activity	Mitochondrial repair and synthesis	MacKenzie and McIntosh (1999) Plant Cell 11: 571-86
Metabolic activity	Commencement of respiration	Debeaujon et al. (2000) Plant Physiol 122: 403-4132
	Water Uptake	Debeaujon et al. (2000) Plant Physiol 122: 403-4132

Other biological activities that are modulated by the imbibition-activated polynucleotides and polynucleotide products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Table below as well as in the Domain section of the Reference Table.

D.2.d. USE OF IMBIBITION AND GERMINATION GENES TO MODULATE THE  
TRANSCRIPTION LEVELS OF OTHER GENES

The expression of many genes is "upregulated" or "downregulated" during imbibition and germination because some imbibition and germination genes are integrated into complex networks that regulate transcription of many other genes. Some imbibition and germination genes are therefore useful for regulating other genes and hence complex phenotypes.

Imbibition-activated polynucleotides may also be differentially transcribed in response to fluctuating developmental-specific polynucleotide levels or concentrations, whether internal or external to a cell, at different times during the plant life cycle to promote associated biological activities. These activities are, by necessity, a small subset of the genes involved in the development process. Furthermore, because development is a continuous process with few clear demarcations between stages, the associated metabolic and biochemical pathways overlap. Some of the changes in gene transcription are summarized in the Table below:

DEVELOPMENTAL PROCESS REGULATED BY IMBIBITION- GERMINATION GENES	PHYSIOLOGICAL/METABOLIC CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL REGULATORY ACTIVITIES ASSOCIATED WITH IMBIBITION AND GERMINATION
Tissue Specialization  - Cotyledon Expansion - Endosperm (???) - Activation of the Shoot Apical Meristem - Activation of the Root Apical Meristem	- Lipid Catabolism - Lipoxygenase Localization - Starch Catabolism - Seed Protein Catabolism - Growth Regulator Production, Transport, Perception, Signaling, Response (e.g., Gibberellins, Ethylene,	- Transcription Factors - Transporters - Kinases - Changes in cytoskeletal protein activity modulating cell structure - Stability of factors for protein translation - Changes in cell



<ul style="list-style-type: none"> <li>- Radicle Growth</li> <li>- Vascular System Development</li> </ul>	<p>Auxin)</p> <ul style="list-style-type: none"> <li>- Global Gene Activation</li> <li>- Transcription Initiation</li> <li>- Sucrose Synthesis and Partitioning</li> <li>- Sucrose catabolism</li> <li>- Sucrose Signaling</li> </ul>	<ul style="list-style-type: none"> <li>wall/membrane structure</li> <li>- Chromatin structure and/or DNA topology</li> <li>- Biosynthetic enzymes</li> <li>- Metabolic enzymes</li> </ul>
<p>Organelle Differentiation and Development</p>	<ul style="list-style-type: none"> <li>- Cell Wall Biosynthesis</li> <li>- Activators of Metabolic Pathways</li> <li>- Protein Remodeling</li> <li>- Cell Wall Biosynthesis</li> <li>- Membrane Repair and Synthesis</li> <li>- Specific Gene Transcription Initiation</li> <li>- Sucrose Mobilization and Partitioning</li> <li>- Sucrose Signaling</li> <li>- Activators of Metabolic Pathways</li> <li>- Auxin Perception, Response and Signaling</li> <li>- Protein Remodeling</li> <li>- Lipid Mobilization, Metabolism and Biosynthesis</li> <li>- Protein Transport, Metabolism, and Biosynthesis</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> <li>- Changes in cytoskeletal protein activity modulating cell structure</li> <li>- Stability of factors for protein translation</li> <li>- Changes in cell wall/membrane structure</li> <li>- Chromatin structure and/or DNA topology</li> <li>- Biosynthetic enzymes</li> <li>- Metabolic enzymes</li> </ul>

DNA Repair	<ul style="list-style-type: none"> <li>- Cell Division</li> <li>- Cell Cycle Control</li> <li>- DNA Replication</li> <li>- Specific Gene Transcription Initiation</li> <li>- Protein Remodeling</li> <li>- Protein Synthesis</li> <li>- Repressors of Senescence</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones for protein translation</li> <li>- Changes in cell wall/membrane structure</li> <li>- Chromatin structure and/or DNA topology</li> <li>- Biosynthetic enzymes</li> </ul>
Cellular Metabolism	<ul style="list-style-type: none"> <li>- Lipid Catabolism                             <ul style="list-style-type: none"> <li>- oxidation</li> <li>- Glyoxylate cycle</li> <li>- Citric acid cycle</li> <li>- Gluconeogenesis</li> </ul> </li> <li>- Sucrose Synthesis and Partitioning</li> <li>- Starch Catabolism</li> <li>- Seed Protein Catabolism                             <ul style="list-style-type: none"> <li>- Asparagine Synthesis and Transport</li> </ul> </li> <li>- Sucrose catabolism</li> <li>- Sucrose Signaling</li> <li>- Ribosome/polysome production and maintenance</li> <li>- Housekeeping genes</li> <li>- Respiration</li> <li>- Photosynthesis</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Kinases</li> <li>- Chaperones</li> <li>- Translation Initiation Factors</li> <li>- Biosynthetic Enzymes</li> <li>- Metabolic Enzymes</li> </ul>

Changes in the processes of germination are the result of modulation of the activities of one

or more of these many germination genes and gene products. These genes and/or products are responsible for effects on traits such as fast germination, plant vigor and seed yield, especially when plants are growing in the presence of biotic or abiotic stresses or when they are growing in barren conditions or soils depleted of certain minerals.

5 Germination genes and gene products can act alone or in combination as described in the introduction. Of particular interest are combination of these genes and gene products with those that modulate stress tolerance and/or metabolism. Stress tolerance and metabolism genes and gene products are described in more detail in the sections below.

#### Use of Promoters of Imbibition and Germination Genes

10 These promoters can be used to control expression of any polynucleotide, plant or non-plant, in a plant host. Selected promoters when operably linked to a coding sequence can direct synthesis of the protein in specific cell types or to loss of a protein product, for example when the coding sequence is in the antisense configuration. They are thus useful in controlling changes in imbibition and germination phenotypes or enabling novel proteins to be made in germinating seeds.

I.D.2. EARLY SEEDLING-PHASE SPECIFIC RESPONSIVE GENES, GENE  
COMPONENTS AND PRODUCTS

One of the more active stages of the plant life cycle is a few days after germination is  
complete, also referred to as the early seedling phase. During this period the plant begins  
development and growth of the first leaves, roots, and other organs not found in the embryo.  
Generally this stage begins when germination ends. The first sign that germination has been  
completed is usually that there is an increase in length and fresh weight of the radicle.

a) IDENTIFICATION OF EARLY SEEDLING PHASE GENES,  
GENE COMPONENTS AND PRODUCTS

These genes defined and identified herein are capable of modulating one or more  
processes of development and growth of many plant organs as described below. These genes  
and gene products can regulate a number of plant traits to modulate yield. Examples of such  
early seedling phase genes and gene products are shown in the Reference and Sequence, Knock-  
in, Knock-out and MA-diff Tables. The functions of the protein of some of these genes are also  
given in these Tables.

Early Seedling Genes Identified by Phenotypic Observations

Some early seedling genes were discovered and characterized from a much larger set of  
genes by experiments designed to find genes that cause phenotypic changes in germinating seeds  
as the transitioned into seedlings.

In these experiments, leaf genes were identified by either (1) ectopic expression of a  
cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then cultivated and  
one or more of the following leaf phenotypes, which varied from the parental "wild-type", were  
observed:

- Abnormal growth
- Abnormal cotyledons or root growth
  - Reduced growth
- Abnormal first leaf
- Abnormal hypocotyl

- Abnormal pigmentation

The genes identified by these phenotypes are given in the Knock-in and Knock-out Tables.

#### Early Seedling Phase Genes Identified By Differential Expression

Such genes are active or potentially active to a greater extent in developing and rapidly growing cells, tissues and organs, as exemplified by development and growth of a seedling 3 or 4 days after planting a seed. These genes herein were also discovered and characterized from a much larger set of genes in experiments designed to find genes. Early seedling phase genes were identified by measuring the relative levels of mRNA products in a seedling 3 or 4 days after planting a seed versus a sterilized seed. Specifically, mRNA was isolated from aerial portion of a seedling 3 or 4 days after planting a seed and compared to mRNA isolated from a sterilized seed utilizing microarray procedures. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Sqn (relating to SMD 7133, SMD 7137)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Early Seedling Phase genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Early Seedling Phase Genes Identified By Cluster Analyses Of Differential Expression

#### Early Seedling Phase Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Early Seedling Phase genes is any group in the MA\_clust

that comprises a cDNA ID that also appears in Expt ID Squ (relating to SMD 7133, SMD 7137) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

5      Early Seedling Phase Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Early Seedling Phase genes. A group in the MA\_clust is considered a Early Seedling Phase pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

10      Early Seedling Phase Genes Identified By Amino Acid Sequence Similarity

Early Seedling Phase genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Early Seedling Phase genes. Groups of Early Seedling Phase genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Early Seedling Phase pathway or network is a group of proteins that also exhibits Early Seedling Phase functions/utilities.

Of particular interest are early seedling phase genes that are differentially expressed 3 or 4 days after planting a seed but not differentially expressed germinating seeds and/or mature leaves.

20      Examples of phenotypes, biochemical activities, and transcription profiles that can be modulated by these genes and gene products are described above and below.

25      D.5.b. USE OF EARLY SEEDLING GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE PHENOTYPES

Rapid, efficient establishment of a seedling is very important in commercial agriculture and horticulture. It is also vital that resources are approximately partitioned between shoot and root to facilitate adaptive growth. Phototropism and geotropism need to be established. All these require post-germination process to be sustained to ensure that vigorous seedlings are produced. Early seedling phase genes, gene components and products are useful to manipulate these and other processes.

I. Development

The early seedling phase genes, gene components and products of the instant invention are useful to modulate one or more processes of the stages of leaf morphogenesis including: stage 1- organogenesis that gives rise to the leaf primordium; stage 2- delimiting basic morphological domains; and stage 3- a coordinated processes of cell division, expansion, and differentiation. Early seedling phase genes include those genes that terminate as well as initiate leaf development. Modulating any or all of the processes leads to beneficial effects at specific locations .

A. Gene Sequences Affecting Types of Leaves

Applicants provide with these genes, gene components and gene products the means to modulate one or more of the following types of leaves, and stem:

5. Cotyledons
6. Major Leaves

B. Gene Sequences Affecting Cell properties

These genes, gene components and gene products are useful to modulate changes in:

10. Cell size
11. Cell division, rate and direction
12. Cell elongation
13. Cell differentiation
14. Xylem and phloem cell numbers
15. Cell wall composition
16. All cell types

C. Gene Sequences Affecting Leaf Architecture:

The following properties of a leaf are useful to modulate to change overall leaf architecture:

4. Veination – Improvements in photosynthetic efficiency, stress tolerance efficiency of solute and nutrient movement to and from the leaf are accomplished by increases or decreases in:
  - (c) Vein placement

- (d) Number of cells in the vein
- 5. Shape
  - (c) Elongated versus rounded
  - (d) Symmetry, around either
    - abaxial-adaxial (dorsiventral) axis
    - apical-basal (proximodistal) axis
    - margin-blade-midrib (lateral) axis

#### H. Genes Sequences Influencing Leaf Responses

Shoot apical meristem cells differentiate to become leaf primordia that eventually develop into leaves. The genes, gene components and gene products of this invention are useful to modulate any one or all of these growth and development processes, by affecting timing and rate or planes of cell divisions for example, in response to the internal plant stimuli and/or programs listed below:

- 8. Embryogenesis
- 9. Germination
- 10. Hormones
  - (b) Auxin (for more details see the section on auxin-responsive genes)
- 11. Phototropism
- 12. Coordination of leaf growth and development with that of other organs
  - (f) Roots
  - (g) Stems
- 13. Stress-related programs

#### II. Interaction with the Environment

Successful seedling establishment demands successful interaction with the environment in the soil. Early vegetation genes orchestrate and respond to interactions with the environment. Thus early seedling phase genes are useful for improving interactions between a plant and the environment that include:

- N. Pigment accumulation (see the section on Viability genes for more detail)
- O. Oxygen gain/loss control
- P. Carbon dioxide gain/loss control



- Q. Water gain/loss control
- R. Nutrient transport
- S. Light harvesting
- T. Chloroplast biogenesis
- U. Circadian rhythm control
- V. Light/dark adaptation
- W. Defense systems against biotic and abiotic stresses
- X. Metabolite accumulation
- Y. Secondary metabolite production

### III. Organizing Tissues for Photosynthesis and Metabolism

Following germination and utilization of seed reserves, plant tissues prepare for photosynthesis and seedling metabolism. Leaf meristems, and root meristems participate in these changes before cell differentiation. Many of the uses for plants depend on the success of leaves as the powerhouses for plant growth, their ability to withstand stresses and their chemical composition. Leaves are organs with many different cell types and structures. Most genes of a plant are active in leaves and therefore leaves have very diverse of pathways and physiological processes. Examples of such pathways and processes that are modulated by early seedling phase genes, gene components and products include:

- X. Photosynthesis
- Y. Sugar metabolism
- Z. Starch synthesis
- AA. Starch degradation
- BB. Nitrate and ammonia metabolism
- CC. Amino acid biosynthesis, transport
- DD. Protein biosynthesis
- EE. DNA replication, repair
- FF. Lipid biosynthesis and breakdown
- GG. Protein biosynthesis, storage and breakdown
- HH. Nucleotide transport and metabolism
- II. Cell envelope biogenesis

- JJ. Membrane formation
- KK. Mitochondrial and chloroplast biogenesis
- LL. Transcription and RNA metabolism
- MM. Vitamin biosynthesis
- NN. Steroid and terpenoid biosynthesis
- OO. Devise secondary metabolite synthesis
- PP. Co-enzyme metabolism
- QQ. Flavonoid biosynthesis and degradation
- RR. Synthesis of waxes
- SS. Glyoxylate metabolism
- TT. Hormone perception and response pathways

#### Uses of Plants that Are Modified as Described above

Altering leaf genes or gene products in a plant modifies one or more the following plant traits, to make the plants more useful for specific purposes in agriculture, horticulture and for the production of valuable molecules. The useful plants have at least one of the following:

- A. More seedling vigor
- B. A higher yield of early leaves and their molecular constituents due to different:
  - 6. Number, size, weight, harvest index
  - 7. Composition including and amounts and types of carbohydrates, proteins, oils, waxes, etc.
  - 8. Photosynthetic efficiency e.g. reduced photorespiration
  - 9. Absorption of water and nutrients to enhance yields, including under stresses such as high light, herbicides, and heat.
  - 10. Pathways to accumulate new valuable molecules.
- I. More optimal leaf shape and architecture in early seedling— enhancing photosynthesis and enhancing appeal in ornamental species
  - (e) size
  - (f) number
  - (g) pigment

- J. A better overall plant architecture – enhancing photosynthesis and enhancing appeal in ornamental species
- K. Reduced negative effects of high planting density, by altering leaf placement to be more vertical instead of parallel to the ground, for instance
- L. Better stress tolerance, including without limitation
3. Drought resistance, by decreasing water loss, for example
4. Pathogen resistance
- M. Better overall yield and vigor
- Plant yield of biomass and of constituent molecules and plant vigor are modulated to create benefits by genetically changing:
3. Growth rate of
- (h) Seedling
- (i) Coleoptile elongation
- (j) Young leaves

To change any of the phenotype(s) above, activities of one or more of the early seedling phase genes or gene products are modulated in an organism and the consequence evaluated by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels are altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (Methods. Mol. Biol. 82:259-266 (1998)) with leaf gene constructs and/or screened for variants as in Winkler et al., Plant Physiol. 118: 743-50 (1998) and visually inspected for the desired phenotype and metabolically and/or functionally assayed for altered levels of relevant molecules.

#### D.5.c. USE OF EARLY SEEDLING PHASE GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

Seedlings are complex and their structure, function and properties result from the integration of many processes and biochemical activities. Some of these are known from the published literature and some can be deduced from the genes and their products described in this application. Early seedling phase genes, and gene components are used singly or in combination

to modify these processes and biochemical activities and hence modify the phenotypic and trait characteristics described above. Examples of the processes and metabolic activities are given in the Table below. The resulting changes are measured according to the citations included in the Table.

5

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Metabolism – anabolic and catabolic	G. Farnesyl ation	Pei et al., <u>Science</u> 282: 287-290 (1998); Cutler et al., <u>Science</u> 273: 1239 (1996)
	H. Cell Wall Biosynt hesis	Goupil et al., <u>J Exptl. Botany</u> 49:1855-62 (1998)
	I. Nitrogen Metabol ism	Walch-Liu et al., <u>J Exppt. Botany</u> 51, 227-237 (2000)
	J. Seconda ry Metabol ite Biosynt hesis and Degrada tion	
Water Conservation And Resistance To Drought	A. Production of polyols	Allen et al., <u>Plant Cell</u> 11: 1785- 1798 (1999)

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
And Other Related Stresses	B. Regulation of salt concentration  C. ABA response(s)	Li et al., <u>Science</u> <u>287</u> : 300-303 (2000) Burnett et al., <u>J Exptl. Botany</u> <u>51</u> : 197-205 (2000) Raschke, In: <u>Stomatal Function</u> , Zeiger et al. Eds., 253-279 (1987)
Transport Anion and Cation Fluxes	(i) Ca <sup>2+</sup> Accumulation  (a) K <sup>+</sup> Fluxes (b) Na <sup>+</sup> Fluxes 1. Receptor – ligand binding 2. Anion and Cation fluxes	Lacombe et al., <u>Plant Cell</u> <u>12</u> : 837- 51 (2000); Wang et al., <u>Plant Physiol.</u> <u>118</u> :1421-1429 (1998); Shi et al., <u>Plant Cell</u> <u>11</u> : 2393- 2406 (1999) Gaymard et al., <u>Cell</u> <u>94</u> :647-655 (1998) Jonak et al., <u>Proc. Natl. Acad. Sci.</u> <u>93</u> : 11274-79 (1996); Sheen, <u>Proc. Natl. Acad. Sci.</u> <u>95</u> : 975-80 (1998); Allen et al., <u>Plant Cell</u> <u>11</u> : 1785-98 (1999)
Carbon Fixation	3. Calvin Cycle 5. Photorespiration n 6. Oxygen	Wingler et al., <u>Philo Trans R Soc Lond B Biol Sci</u> <u>355</u> , 1517-1529 (2000);

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	<p>evolution</p> <p>7. RuBisCO</p> <p>4. Chlorophyll metabolism</p> <p>(ii) Chloropl ast Biogene sis and Metabol ism</p> <p>5. Fatty Acid and Lipid Biosynthesis</p> <p>(iii) Glyoxyl ate metaboli sm</p> <p>(iv) Sugar Transpo rt</p> <p>(v) Starch Biosynt hesis and Degrada</p>	<p>Palecanda et al., <u>Plant Mol Biol</u> <b>46</b>, 89-97 (2001);</p> <p>Baker et al., <u>J Exp Bot</u> <b>52</b>, 615-621 (2001)</p> <p>Chen et al., <u>Acta Biochim Pol</u> <b>41</b>, 447-457 (1999)</p> <p>Imlau et al., <u>PlantCell II</u>, 309-322 (1999)</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	tion	
Hormone Perception and Growth	<p>(vi) Hormone Receptors and Downstream Pathways for</p> <p>(a) ethylene</p> <p>(b) jasmonic acid</p> <p>(c) brassinosteroid</p> <p>(d) gibberellin</p> <p>(e) auxin</p> <p>(f) cytokinin</p> <p>- Activation Of Specific Kinases And Phosphatases</p>	<p>Tieman et al., <u>Plant J</u> <u>26</u>, 47-58 (2001)</p> <p>Hilpert et al., <u>Plant J</u> <u>26</u>, 435-446 (2001)</p> <p>Wenzel et al., <u>Plant Phys</u> <u>124</u>, 813-822 (2000)</p> <p>Dengler and Kang, <u>Curr Opin Plant Biol</u> <u>4</u>, 50-56 (2001)</p> <p>Tantikanjana et al., <u>Genes Dev</u> <u>15</u>, 1577-1580 (2001)</p>
See Imbibition, Shoot Apical Meristem, Root and Leaf sections for more details		

Other biological activities that are modulated by the leaf genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table, for example.

D.5.d. USE OF EARLY SEEDLING PHASE GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE TRANSCRIPTION LEVELS

The expression of many genes is "up regulated" or down regulated" in plants because  
5 some genes and their products are integrated into complex networks that regulate transcription of  
many other genes. Some early seedling phase genes, gene components and products are  
therefore useful for modifying the transcription of other genes and hence complex phenotypes, as  
described above. Profiles of leaf gene activities are described in the Table below with associated  
biological activities. "Up-regulated" profiles are those where the mRNA transcript levels are  
10 higher in young seedlings as compared to the sterilized seeds. "Down-regulated" profiles  
represent higher transcript levels in the plantlet as compared to sterilized seed only.

I.D.3. SIZE AND STATURE GENES, GENE COMPONENTS AND  
PRODUCTS

Great agronomic value can result from modulating the size of a plant as a whole or of any of  
15 its organs. For example, the green revolution came about as a result of creating dwarf wheat plants,  
which produced a higher seed yield than taller plants because they could withstand higher levels and  
inputs of fertilizer and water. Size and stature genes elucidated here are capable of modifying the  
growth of either an organism as a whole or of localized organs or cells. Manipulation of such  
20 genes, gene components and products can enhance many traits of economic interest from increased  
seed and fruit size to increased lodging resistance. Many kinds of genes control the height attained  
by a plant and the size of the organs. For genes additional to the ones in this section other sections  
of the Application should be consulted.

a) Identification of Size and Stature Genes, Gene Components and  
Products

Size and stature genes identified herein are defined as genes, gene components and  
products capable of modulating one or more processes in growth and development, to produce  
changes in size of one or more organs. Examples of such stature genes and gene products are  
25 shown in the Reference, Sequence, Protein Group, Protein Group Matrix, Knock-in, Knock-out,  
30



MA-diff and MA-clust. The biochemical functions of the protein products of many of these genes determined from comparisons with known proteins are also given in the Reference tables.

b) Size And Stature Genes, Gene Components And Products  
Identified By Phenotypic Observations

Mutant plants exhibiting increased or decreased stature in comparison to parental wild-type plants were used to identify size and stature genes. In these experiments, size and stature genes were identified by either (1) the ectopic expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then cultivated and stature genes were identified from plants that were smaller than the parental "wild-type". The phenotypes and gene mutations associated with them are given in Tables

Examples of phenotypes, biochemical activities, or transcript profiles that are modulated using these genes are described above and below.

c) Use Of Size And Stature Genes, Gene Components And Products  
To Modulate Phenotypes

Typically, these genes can cause or regulate cell division, rate and time; and also cell size and shape. Many produce their effects via meristems. These genes can be divided into three classes. One class of genes acts during cytokinesis and/or karyokinesis, such as mitosis and/or meiosis. A second class is involved in cell growth; examples include genes regulating metabolism and nutrient uptake pathways. Another class includes genes that control pathways that regulate or constrain cell division and growth. Examples of these pathways include those specifying hormone biosynthesis, hormone sensing and pathways activated by hormones.

Size and stature genes and gene components are useful to selectively alter the size of organs and stems and so make plants specifically improved for agriculture, horticulture and other industries. There are a huge number of utilities. For example, reductions in height of specific ornamentals, crops and tree species can be beneficial, while increasing height of others may be beneficial.

Increasing the length of the floral stems of cut flowers in some species would be useful, while increasing leaf size in others would be economically attractive. Enhancing the size of

specific plant parts, such as seeds, to enhance yields by stimulating hormone (Brassinolide) synthesis specifically in these cells would be beneficial. Another application would be to stimulate early flowering by altering levels of gibberellic acid in specific cells. Changes in organ size and biomass also results in changes in the mass of constituent molecules. This makes the utilities of size and stature genes useful for the production of valuable molecules in parts of plants, for extraction by the chemical and pharmaceutical industries.

Examples of phenotypes that can be modulated by the genes and gene components are described above and below:

I. Cellular Level:

Size and stature genes and gene products can be used to modulate cellular changes in:

- A. Cell size
- B. Cell shape
- C. Cell division, rate and direction
- D. Cell elongation
- E. Cell differentiation
- F. Stomata number
- G. Trichome number

II. Organ Level

The genes of the invention are useful to regulate the development and growth of:

A. Roots

- 1. Primary
- 2. Lateral
- 3. Root hairs
- 4. Root cap
- 5. Apical meristem
- 6. Epidermis
- 7. Cortex
- 8. Stele

B. Stem

1. Phloem
2. Xylem
3. Nodes
4. Internodes
5. Shoot apical meristem

C. Leaves

1. Cauline
2. Rosette
3. Petioles

D. Flowers

1. Receptacle
2. Sepals, Petals, and Tepals
  - (a) Color
  - (b) Shape
  - (c) Size
  - (d) Number
  - (e) Petal Drop
3. Androecium
4. Stamen
5. Anther
6. Pollen
7. Sterility
8. Size, shape, weight, color
9. Filament
10. Gynoecium
11. Carpel
12. Ovary
13. Style
14. Stigma
15. Ovule
  - (a) Size, shape, number

- 16. Pedicel and Peduncle
- 17. Flowering time
- 18. Fertilization
- E. Seeds
  - 1. Placenta
  - 2. Embryo
  - 3. Cotyledon
  - 4. Endosperm
  - 5. Suspensor
  - 6. Seed coat (testa)
- F. Fruits
  - 1. Pericarp – thickness, texture
  - 2. Exocarp
  - 3. Mesocarp
  - 4. Endocarp

### III. Overall Organism Level

The following traits can be modulated with the genes and gene products of this invention to affect the traits of a plant as a whole:

- A. Architecture
  - 1. Branching
  - 2. Ornamental architecture
  - 3. Shade avoidance
  - 4. Planting density effects
  - 5. Wind resistance
- B. Vigor
  - 1. Increased biomass
  - 2. Drought tolerance

To regulate any of the phenotype(s) above, activities of one or more of the sizing genes or gene products are modulated in an organism and tested by screening for the desired trait.

Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (*Methods. Mol. Biol.* 82:259-266 (1998)) and/or screened for variants as in Winkler et al., (*Plant Physiol.* 118: 743-50 1998) and visually inspected for the desired phenotype or metabolically and/or functionally assayed.

D.3.b. USE OF SIZE AND STATURE GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

Many metabolic and developmental processes can be modulated by size and stature genes and gene components to achieve the phenotypic characteristics exemplified above. Some of these are listed below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth and Development	Gibberellic Acid Biosynthesis Gibberellic Acid Receptor and Downstream Pathways	Swain SM, Tseng Ts, Olszewski NE. Altered expression of spindly affects gibberellin response and plant development. <i>Plant Physiol</i> 2001 Jul;126(3):1174-85  Hooley, R. Gibberellins: perception, transduction, and responses. <i>Plant Mol. Biol.</i> 1994 26:1529-1555.  Hooley, R. Gibberellins: perception, transduction, and responses. <i>Plant Mol. Biol.</i> 1994 26:1529-1555.  Perata, P, Matsukura, C, Vernieri, P, Yamaguchi, J, Sugar repression of a gibberellin-dependent signaling pathway in barley embryos. <i>Plant Cell</i> 1997

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Brassinolide Biosynthesis Brassinolide Receptors, Degradation of Brassinolide Pathways affected by Brassinolide	9:2197-2208.  Noguchi T, Fujioka S, Choe S, Takatsuto S, Tax FE, Yoshida S, Feldmann KA. Biosynthetic pathways of brassinolide in Arabidopsis. Plant Physiol 2000 Sep;124(1):201-9  Wang ZY, Seto H, Fujioka S, Yoshida S, Chory J. BRI1 is a critical component of a plasma-membrane receptor for plant steroids. Nature 2001 Mar 15;410(6826):380-3  Neff MM, Nguyen SM, Malanchruvil EJ, Fujioka S, Noguchi T, Seto H, Tsubuki M, Honda T, Takatsuto S, Yoshida S, Chory J. BAS1: A gene regulating brassinosteroid levels and light responsiveness in Arabidopsis. Proc Natl Acad Sci U S A 1999 Dec 21;96(26):15316-23  Kang JG, Yun J, Kim DH, Chung KS, Fujioka S, Kim JI, Dae HW, Yoshida S, Takatsuto S, Song PS, Park CM. Light and brassinosteroid signals are integrated via a dark-induced small G protein in etiolated seedling growth. Cell 2001 Jun 1;105(5):625-36
	Cytokinin biosynthesis Cytokinin receptor Degradation of Cytokinin Pathways affected by Cytokinin	Mok DW, Mok MC. Cytokinin metabolism and action. Annu Rev Plant Physiol Plant Mol Biol 2001;52:89-118  Schmulling T. CREAm of

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	<p>Auxin Biosynthesis Auxin Receptor Auxin Degradation Pathways affected by Auxins Auxin transport</p>	<p>cytokinin signalling: receptor identified. Trends Plant Sci 2001 Jul;6(7):281-4</p> <p>Mok DW, Mok MC. Cytokinin metabolism and action. Annu Rev Plant Physiol Plant Mol Biol 2001;52:89-118</p> <p>Seyedi M, Selstam E, Timko MP, Sundqvist C. The cytokinin 2-isopentenyladenine causes partial reversion to skotomorphogenesis and induces formation of prolamellar bodies and protochlorophyllide657 in the lip1 mutant of pea. Physiol Plant 2001 Jun;112(2):261-272</p> <p>Zhao Y, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, Weigel D, Chory J. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. Science 2001 Jan 12;291(5502):306-9</p> <p>Abel S, Ballas N, Wong LM, Theologis A. DNA elements responsive to auxin. Bioessays 1996 Aug;18(8):647-54</p> <p>del Pozo JC, Estelle M. Function of the ubiquitin-proteasome pathway in auxin response. Trends Plant Sci 1999 Mar;4(3):107-112.</p> <p>Rahman A, Amakawa T, Goto N, Tsurumi S. Auxin is a positive regulator for ethylene-mediated response in the</p>

0093525 082101

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		<p>growth of Arabidopsis roots. Plant Cell Physiol 2001 Mar; 42(3):301-7</p> <p>Zhao Y, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, Weigel D, Chory J. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. Science 2001 Jan 12;291(5502):306-9</p> <p>Abel S, Ballas N, Wong LM, Theologis A. DNA elements responsive to auxin. Bioessays 1996 Aug;18(8):647-54</p> <p>del Pozo JC, Estelle M. Function of the ubiquitin-proteasome pathway in auxin response. Trends Plant Sci 1999 Mar;4(3):107-112.</p> <p>Rahman A, Amakawa T, Goto N, Tsurumi S. Auxin is a positive regulator for ethylene-mediated response in the growth of Arabidopsis roots. Plant Cell Physiol 2001 Mar; 42(3):301-7</p> <p>Gil P, Dewey E, Friml J, Zhao Y, Snowden KC, Putterill J, Palme K, Estelle M, Chory J. BIG: a calossin-like protein required for polar auxin transport in Arabidopsis. Genes Dev. 2001 Aug 1;15(15):1985-97</p> <p>Estelle M., Polar auxin transport. New support for an old model. Plant Cell 1998 Nov;10(11):1775-8</p>



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Cell wall growth	Cosgrove DJ., Loosening of plant cell walls by expansins. Nature 2000 Sep 21;407(6802):321-6

Other biological activities that are modulated by the stature genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table, for example.

Changes in the size, vigor, or yield of a plant are the result of modulation of the activities of one or more of these many size and stature genes and gene products. While size and stature polynucleotides and gene products can act alone, combinations of these polynucleotides and also with others that also affect growth and development are especially useful.

#### Use of Promoters of "Size and Stature" Genes

Promoters of "size and stature" genes are useful for controlling the transcription of any desired polynucleotides, both plant and non-plant. They can be discovered from the "size and stature" genes in the Reference Tables, and their patterns of activity from the MA Tables. When operably linked to any polynucleotide encoding a protein, and inserted into a plant, the protein will be synthesized in those cells in which the promoter is active. Many "size and stature" genes will function in meristems, so the promoters will be useful for expressing proteins in meristems. The promoters can be used to cause loss of, as well as synthesis of, specific proteins via antisense and sense suppression approaches.

I.D.4. SHOOT-APICAL MERISTEM GENES, GENE COMPONENTS AND  
PRODUCTS

5 New organs, stems, leaves, branches and inflorescences develop from the stem apical  
meristem (SAM). The growth structure and architecture of the plant therefore depends on the  
behavior of SAMs. Shoot apical meristems (SAMs) are comprised of a number of morphologically  
undifferentiated, dividing cells located at the tips of shoots. SAM genes elucidated here are capable  
of modifying the activity of SAMs and thereby many traits of economic interest from ornamental  
leaf shape to organ number to responses to plant density.

a) IDENTIFICATION OF SAM GENES, GENE COMPONENTS  
AND PRODUCTS

10  
15 SAM genes identified herein are defined as genes, gene components and products capable of  
modulating one or more processes or functions of SAMs as described below. Regulation of SAM  
genes and gene products are useful to control many plant traits including architecture, yield and  
vigor. Examples of such SAM genes and gene products are shown in the Reference, Sequence,  
Protein Group, Protein Group Matrix, phenotype and MA-diff Tables. The functions of many of the  
protein products of these genes are also given in the Reference tables.

20 Sam Genes, Gene Components And Products Identified By Phenotypic Observations

25 SAM genes were discovered and characterized from a much larger set of genes by  
experiments designed to find genes that cause phenotypic changes in leaf morphology, such as  
cotyledon or leaf fusion. In these experiments, SAM genes were identified by either (1) ectopic  
expression of a cDNA in a plant or (2) mutagenesis of the plant genome. The plants were then  
cultivated and one or more of the following phenotypes, which varied from the parental "wild-  
type", was observed:

- I. Cotyledon  
- Fused  
II. Leaves  
- Fused

- Leaf placement on stems
- III. Branching
  - Number
- IV. Flowers
  - Petals fused
  - Altered bolting
  - Early bolting
  - Late bolting
  - Strong bolting
  - Weak bolting
  - Abnormal branching

For more experimental detail see the Example section below. The genes identified by these results of the phenotypes that are shown in Knock-in and Knock-out Tables.

Sam Genes, Gene Components And Products Identified By Differential Expression

SAM genes were also identified in experiments designed to find genes whose mRNA products are associated specifically or preferentially with SAMs. The concentration of mRNA products in the arabidopsis plant with the SHOOTMERISTEMLESS (STM) gene knocked-out was measured relative to the concentration in the parental, non-mutant plant. The Arabidopsis STM gene is required for embryonic SAM formation. The STM gene encodes a Knotted1 (Kn1) type of homeodomain protein. Homeodomain proteins regulate transcription of many genes in many species and have been shown to play a role in the regulation of translation as well. Seedlings homozygous for recessive loss-of-function alleles germinate with roots, a hypocotyl, and cotyledons, but no SAM is formed. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108478, 108479, 108480, 108481, 108598, 108535, 108536, 108435). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Meristem genes are those sequences that showed differential expression as compared to

controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

Meristem Genes Identified By Cluster Analyses Of Differential Expression

Meristem Genes Identified By Correlation To Genes That Are Differentially

Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Meristem genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108478, 108479, 108480, 108481, 108598, 108535, 108536, 108435 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Meristem Genes Identified By Correlation To Genes That Cause

Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Meristem genes. A group in the MA\_clust is considered a Meristem pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Meristem Genes Identified By Amino Acid Sequence Similarity

Meristem genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Meristem genes. Groups of Meristem genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Meristem pathway or network is a group of proteins that also exhibits Meristem functions/utilities.

Examples of phenotypes, biochemical activities, and transcription profiles that can be modulated by SAM genes and gene products are described above and below.

b) USE OF SAM GENES, GENE COMPONENTS AND PRODUCTS  
TO MODULATE PHENOTYPES

With the SAM genes and gene products of the invention, Applicants provide the means to modulate one or more of the following types of SAMs:

1. Embryonic meristem
2. Vegetative lateral SAMs
3. Inflorescence lateral SAMs
4. Floral meristems
5. Adventitious SAM

The SAM genes of the instant invention are useful for modulating one or more processes of SAM structure and/or function including (I) cell size and division; (II) cell differentiation and organ primordia.

I. Cell Size and Division

A. Cell Properties

SAM genes and gene products can be used to modulate changes in:

1. Cell size
2. Cell division, rate and direction
3. Cell division symmetry

A key attribute of the SAM is its capacity for self-renewal. The self-renewing initial cell population resides in the central zone of the SAM. A small number of slowly dividing initial cells (typically 2 to 4 per layer) act as a self-replenishing population, whereas some of their descendants, pushed out onto the flanks of the SAM, differentiate into leaves. Other descendants, displaced below the SAM, differentiate into stem. The immediate descendants of the initial cells divide further, amplifying the cell population before being incorporated into leaf or stem primordia.

The genes and gene components of this invention are useful for modulating any one or all of these cell division processes generally, as in timing and rate, for example. In addition, the polynucleotides and polypeptides of the invention can control the response of these processes to the internal plant programs associated with:

1. Embryogenesis

2. Hormone responses  
Cytokinin (inhibitory for root development, see section on cytokinin-responsive genes)
3. Coordination of growth and development with that of other plant organs
  - a. Leaves
  - b. Flowers
  - c. Seeds
  - d. Fruits

SAM genes can also be used to control the response of these processes to changes in the environment, including heat, cold, drought, high light and nutrition.

#### B. Sam Cell Patterns And Organization

Although SAMs appear as small regions of morphological undifferentiated dividing cells, a group of morphologically undifferentiated dividing cells does not necessarily constitute a SAM. Rather, evidence indicates that SAMs are highly organized or patterned regions of the plant in which many important events in early organogenesis occur. Thus, the term "SAM" is used to denote a highly organized structure and site of pattern formation. The invention also permits engineering of specific as well as overall features of SAM architecture as follows:

1. Zones
  - a. Central
  - b. Peripheral
  - c. Rib
2. Layers
  - a. L1
  - b. L2
  - c. L3
3. Symmetry

#### II Cell Differentiation And Organ Primordia

The apical meristem in many species first undergoes a vegetative phase whereby cells set aside from the apex become leaf primordia with an axillary vegetative meristem. Upon floral

induction, the apical meristem is converted to an inflorescence meristem. The inflorescence meristem arises in the axils of modified leaves and is indeterminate, producing whorls or rings of floral organ primordia. In species which produce terminal flowers, the apical meristem is determinate and eventually adopts a third identity, that of a floral meristem. Examples of the plant properties that the genes and gene products of the invention can be used to modulate include:

1. Indeterminancy
  - a. Inhibiting or increasing differentiation
  - b. Enhancing plant growth and yield
2. Symmetry
  - a. Symmetry of organs developed
  - b. Symmetry of arrangement of organs, such as leaves, petals, flowers, etc.
3. Leaf fate and timing (for more detail see leaf section)
  - a. Internode length modulation
    - i. Longer internodes to increase shade avoidance
    - ii. Shorter internodes to favor leaf development
4. Floral fate and timing of flowering (for more detail see the reproduction section)

#### Uses Of Plants Modified As Described Above Using SAM Genes, Gene Components And Products

Because SAMs determine the architecture of the plant, modified plants will be useful in many agricultural, horticultural, forestry and other industrial sectors. Plants with a different shape, numbers of flowers and seed and fruits will have altered yields of plant parts. For example, plants with more branches can produce more flowers, seed or fruits. Trees without lateral branches will produce long lengths of clean timber. Plants with greater yields of specific plant parts will be useful sources of constituent chemicals. Such plants will have, for example:

- A. More prolific leaf development (see Leaf section for more detail)
- B. Better optimized stem and shoot development (see Stem section for more detail)
- C. Adventitious shoots

D. More flowers, seeds, and fruits (see Reproduction section for more detail)

E. Enhanced vigor, including growth rate of:

1. Whole plant, including height, flowering time, etc.
2. Seedling
3. Coleoptile elongation
4. Young leaves
5. Flowers
6. Seeds
7. Fruit

F. Higher yields based on:

1. Biomass  
Fresh and dry weight during any time in plant life, including  
maturation and senescence
2. Number of flowers
3. Seed yield
  - a. Number, size, weight, harvest index
  - b. Content and composition, e.g. amino acid, jasmonate, oil, protein  
and starch
4. Fruit yield
  - a. Number, size, weight, harvest index
  - b. Content and composition, e.g. amino acid, jasmonate, oil, protein  
and starch

To regulate any of the phenotype(s) above, activities of one or more of the SAM genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Dolan et al. (1993, Development 119: 71-84), Dolan et al. (1997, Development 124: 1789-98), Crawford and Glass



(1998, Trends Plant Science 3: 389-95), Wang et al. (1998, PNAS USA 95: 15134-39), Gaxiola et al. (1998, PNAS USA 95: 4046-50), Apse et al. (1999, Science 285: 1256-58), Fisher and Long (1992, Nature 357: 655-60), Schneider et al. (1998, Genes Devel 12: 2013-21) and Hirsch (1999, Curr Opin Plant Biol. 2: 320-326).

5

c) USE OF SAM GENES AND GENE COMPONENTS TO  
MODULATE BIOCHEMICAL ACTIVITIES

SAM genes and gene components are useful for modulating biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation And Development	Leaf shape and inflorescence and flower morphology systems Activities of SAM transcriptional regulatory proteins.	Chuck, G. et al., 1996 Plant Cell 8: 1227-1289. Schneeberger et al., 1998 Development 125: 2857-2865.
	Meristem size and organ number determinants - Regulated by Receptor Kinases - Receptor kinase location and activity.	- Kayes, J.M. and Clark, S.E. 1998 Development 125: 3843- 3851. - Jeong, S. et al., 1999 Plant Cell 11: 1925-1934.
	Meristem proliferation activities	Tantikanjana, T. Genes and Development. June 15, 2001. 15(12):1577-1588.
- Internode elongation	Hormone signaling pathways	Yamamuro, C. et al., 2000 Plant Cell. 12: 1591-1605.
Hormone	Levels of growth hormones	Kusaba, S. et al; 1998 Plant

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Perception	including gibberellic acid, Auxin and cytokinin.	Physiology 116(2):471-476.
	Gibberellic acid biosynthesis GA biosynthetic enzyme GA-20 oxidase is a required step in GA biosynthesis. GA-20 oxidase is Regulated by some SAM gene products.	Modulation of GA perception and function can be assayed as described in Sakamoto, T. et al. 2001 Genes and Development 15: 581-590.
	Over expression of SAM genes can lead to reduced internode elongation, reduced cell elongation and reduced cell expansion.	Sakamoto, T. et al. 2001. Genes and Development 15: 581-590.
	Cytokinin Receptor activity	Inoue, T. et al., Nature 409:1060-1063.
	SAM gene products can affect the activity of Auxin dependent postranscriptional gene protein expression.	Sieberer, T. et al., 2000 Current Biology 10:1595-1598. del Pozo, J. C.; Estelle, M. PNAS (USA) 1999. 96(26):15342-15347.
	SAM gene products can affect Auxin Perception/metabolism in the meristem to produce useful changes in plant architecture.	Tantikanjana, T. Genes and Development. June 15, 2001. 15(12):1577-1588.
Leaf senescence	SAM gene products can increase and decrease leaf senescence	Ori, N. et al; Plant Cell. June, 1999. 11(6):1073-1080.

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	rate. This can be done by modulating cytokinin hormone levels.	
	Cytokinin effect on cell division and expansion.	Beemster, Gerrit T. S.; Baskin, Tobias I. 2000 Plant Physiology 124:1718-1727.
Adventitious shoot formation	Alter growth hormone status.	Kusaba, S. et al; 1998 Plant Physiology 116(2):471-476
	Ectopic expression of SAM genes in leaf or other non SAM organs or tissue can produce shoots	Chuck, G. 1996 Plant Cell 8: 1227-1289.
	Pathways comprising isopentenyl transferase (ipt)	

Other biological activities that can be modulated by the SAM genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

5

d) USE OF SAM GENES, GENE COMPONENTS AND PRODUCTS  
TO MODULATE TRANSCRIPTION LEVELS OF OTHER GENES

The expression of many genes is "upregulated" or "downregulated" in the SAM mutants because some of the SAM genes are integrated into complex networks that regulate the transcription of many other genes. Some SAM genes and gene components are therefore useful for modifying the transcription of other genes and hence complex phenotypes as described above. Profiles of genes altered by SAM mutations and genes are described in the Table below with associated biological activities. "Up-regulated" profiles are for genes whose mRNA levels are higher in the

stm plants as compared to parental wild-type plants; and vice-versa for "down-regulated" profiles.

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING SAM GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
Up Regulated Transcripts	Genes repressed by SAMs directly or indirectly	<ul style="list-style-type: none"> <li>• Altered Auxin/cytokinin hormone ratio and perception.</li> <li>• Increased/decreased cell expansion – promoting effects of brassinosteroids and gibberellic acids, due to altered levels of biosynthetic pathway enzymes and or the amount of functional hormone receptor.</li> <li>• Increased or decreased rate of cell division.</li> <li>• Altered planes of cell division</li> <li>• Increased or decreased rate and extent of cell expansion.</li> <li>• Increased or</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic Enzymes</li> <li>• Cell Membrane Structure</li> <li>• Kinases, Phosphatases, G-Proteins</li> <li>• Transcription Activators/Repressors</li> <li>• Transcription coactivators/corepressor s</li> <li>• Chromatin Structure And/Or Localized DNA Topology Proteins</li> <li>• Cell Wall Proteins</li> <li>• Translational activators/repressors</li> <li>• Cell wall proteins involved in cell rigidity e.g. extensin, glycine rich proteins.</li> <li>• Cell cycle regulatory proteins such as cyclins and cyclin dependent</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING SAM GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
		decreased rigidity of cell ways.	protein kinases (CDKs).
Down-Regulated Transcripts	Genes involved in SAM cells and genes whose expression is induced by SAMs	<ul style="list-style-type: none"> <li>• Altered pattern of organs emerging from the meristem</li> <li>• Increased or decreased the number of cells partitioned into a lateral organ.</li> <li>• Altered apical dominance due to suppression of lateral bud growth.</li> <li>• Altered apical dominance due to releasing of axillary meristems from repression.</li> <li>• Increased/or decreased production of adventitious meristems.</li> <li>• Increased potential to form somatic embryos.</li> </ul>	<ul style="list-style-type: none"> <li>• Auxin transporter proteins</li> <li>• Auxin receptor proteins</li> <li>• Cytokinin receptor proteins</li> <li>• Gibberellic acid receptor proteins</li> <li>• Brassinolide receptor proteins</li> <li>• Hormone biosynthesis proteins</li> <li>• Hormone degradation proteins</li> <li>• Hormone conjugation proteins</li> <li>• Ubiquitin conjugating enzymes.</li> <li>• Receptor kinase signal transduction</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING SAM GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
		<ul style="list-style-type: none"> <li>• Altered cell signaling pathways</li> <li>• Altered hormone levels</li> </ul>	

SAM genes and gene products can be modulated alone or in combination as described in the introduction. Of particular interest are combination of these genes and gene products with those that modulate hormone responsive pathways. Hormone responsive genes and gene products are described in more detail in the sections below.

#### Use Of Sam Gene Promoters To Modify Sams

Promoters of SAM genes, as described in the Reference tables, for example, can be used to modulate transcription of coding sequences in SAM cells to influence growth, differentiation or patterning of development or any of the phenotypes or biological activities above. For example, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as a SAM gene when the desired sequence is operably linked to the promoter of the SAM gene.

A specific instance is linking of a SAM gene promoter normally active in floral meristem primordia, to a phytotoxic protein coding sequence to inhibit apical meristem switching into an inflorescence and/or floral meristem, thereby preventing flowering.

SAM gene promoters can also be used to induce transcription of antisense RNA copies of a gene or an RNA variant to achieve reduced synthesis of a specific protein in specific SAM cells. This provides an alternative way to the example above, to prevent flowering.

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS WITH MODIFIED LEVELS
Up regulated Transcripts	<p>Genes involved in leaf, stem and root cell differentiation, cell division, cell expansion</p> <p>Genes involved in positive regulation of root, stem and leaf genes</p> <p>Repressors of root and other organ cell types e.g. flowers</p>	<ul style="list-style-type: none"> <li>• Leaf cells proliferate and differentiate;</li> <li>• Leaf structures form and expand</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors, signal transduction proteins, kinase and phosphatases</li> <li>• Chromatin remodeling</li> <li>• Hormone biosynthesis enzymes</li> <li>• Receptors</li> </ul>
	Genes involved in photosynthesis	<ul style="list-style-type: none"> <li>• Photosynthesis and plastid differentiation</li> <li>• Calvin cycle activated</li> <li>• Chloroplast biogenesis and plastid differentiation activated</li> </ul>	<ul style="list-style-type: none"> <li>• Light harvesting coupled to ATP production</li> <li>• Chlorophyll biosynthesis</li> <li>• Ribulose Bisphosphate carboxylase</li> <li>• Chloroplast membranes synthesis</li> <li>• Chloroplast ribosome biogenesis</li> </ul>
	Other genes involved in metabolism	<ul style="list-style-type: none"> <li>• Starch biosynthesis</li> <li>• Lipid biosynthesis</li> <li>• Nitrogen metabolism – NO<sub>3</sub> reduced and amino acids made</li> <li>• Secondary metabolites produced</li> </ul>	<ul style="list-style-type: none"> <li>• Starch synthase</li> <li>• Nitrate reductase</li> <li>• Terpenoid biosynthesis</li> <li>• Transcription factors</li> <li>• Transporters</li> <li>• Kinases</li> <li>• Phosphatases and signal transduction protein</li> <li>• Chromatin</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS WITH MODIFIED LEVELS
			structure modulators
Down regulated genes	<p>Genes involved in negative regulation of root, stem and leaf genes</p> <p>Genes involved in other organs e.g. flowers</p>	<ul style="list-style-type: none"> <li>• Leaf genes activated and leaf functions induced</li> <li>• Other organs not induced</li> <li>• Leaf, stem and root metabolic pathways induced</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Signal transduction proteins – kinases and phosphatases</li> <li>• Metabolic enzymes</li> <li>• Chromatin remodeling proteins</li> </ul>

While early seedling phase polynucleotides and gene products are used singly, combinations of these polynucleotides are often better to optimize new growth and development patterns. Useful combinations include different leaf polynucleotides and/or gene products with a hormone responsive polynucleotide. These combinations are useful because of the interactions that exist between hormone-regulated pathways, nutritional pathways and development.

#### Use of Early Seedling Phase Gene Promoters

Promoters of early seedling phase genes are useful for transcription of desired polynucleotides, both plant and non-plant. If the gene is expressed only in the post-germination seedling, or in certain kinds of leaf cells, the promoter is used to drive the synthesis of proteins specifically in those cells. For example, extra copies of carbohydrate transporter cDNAs operably linked to a early seedling phase gene promoter and inserted into a plant increase the “sink” strength of leaves. Similarly, early seedling phase promoters are used to drive transcription of metabolic enzymes that alter the oil, starch, protein, or fiber contents of the seedling. Alternatively, the promoters direct expression of non-plant genes that can, for instance, confer resistance to specific pathogen. Additionally the promoters are used to synthesize an antisense mRNA copy of a gene to inactivate the normal gene expression into protein. The



promoters are used to drive synthesis of sense RNAs to inactivate protein production via RNA interference.

I.D.5. VEGETATIVE-PHASE SPECIFIC RESPONSIVE GENES, GENE  
COMPONENTS AND PRODUCTS

Often growth and yield are limited by the ability of a plant to tolerate stress conditions, including water loss. To combat such conditions, plant cells deploy a battery of responses that are controlled by a phase shift, from so called juvenile to adult. These changes at distinct times involve, for example, cotyledons and leaves, guard cells in stomata, and biochemical activities involved with sugar and nitrogen metabolism. These responses depend on the functioning of an internal clock, that becomes entrained to plant development, and a series of downstream signaling events leading to transcription-independent and transcription-dependent stress responses. These responses involve changes in gene expression.

Manipulation of the activation of one or more genes controlling the phase changes are useful to modulate the biological processes and/or phenotypes listed below. Phase responsive genes and gene products can act alone or in combination. Useful combinations include phase responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants.

Phase responsive genes and gene products can function to either increase or dampen the above phenotypes or activities. Characterization of phase responsive genes was carried out using microarray technology. Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing about 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

10 The sequences of the ESTs showing at least two-fold increases or decreases in a mutant of  
11 *Arabidopsis thaliana*, squint, that appears not to undergo phase changes and appears adult-like  
12 throughout its growth cycle, compared with wild type were identified, compared to the Ceres full  
13 length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. The  
14 MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables reports the  
15 results of this analysis, indicating those Ceres clones which are up or down regulated over  
16 controls, thereby indicating the Ceres clones which represent phase responsive genes. The  
17 MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the  
18 transcript levels of the experiment (see EXPT ID: Sqn (relating to SMD 7133, SMD 7137)). For  
19 transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown  
20 for when transcript levels were reduced in root tips as compared to the control. For more  
21 experimental detail see the Example section below.

22 Phase responsive genes are those sequences that showed differential expression as  
23 compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff  
24 MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Phase Responsive Genes Identified By Cluster Analyses Of Differential Expression

##### Phase Responsive Genes Identified By Correlation To Genes That Are Differentially Expressed

25 As described above, the transcription profiles of genes that act together are well  
26 correlated. Applicants not only have identified the genes that are differentially expressed in the  
27 microarray experiments, but also have identified the genes that act in concert with them. The  
28 MA\_clust table indicates groups of genes that have well correlated transcription profiles and  
29 therefore participate in the same pathway or network.

30 A pathway or network of phase responsive genes is any group in the MA\_clust that  
comprises a cDNA ID that also appears in Expt ID Sqn (relating to SMD 7133, SMD 7137) of  
the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Phase Responsive Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of phase responsive genes. A group in the MA\_clust is considered a phase responsive pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Phase Responsive Genes Identified By Amino Acid Sequence Similarity

Phase responsive genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis phase responsive genes. Groups of phase responsive genes are identified in the Protein Grouping table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a phase responsive pathway or network is a group of proteins that also exhibits Phase responsive functions/utilities.

Further, promoters of phase responsive genes, as described in Reference tables, for example, are useful to modulate transcription that is induced by phase or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the phase responsive genes when the desired sequence is operably linked to a promoter of a phase responsive gene.

#### a) USE OF PHASE RESPONSIVE GENES TO MODULATE PHENOTYPES

Phase responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Timing Phenotypes
- Dormancy
- Germination
- Cotyledon opening
- First leaves
- Juvenile to adult transition
- Bolting
- Flowering

- Pollination
- Fertilization
- Seed development
- Seed set
- Fruit Drop
- Senescence
- Epinasty
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation
- Number of Flowers, Seeds, Branches, and/or Leaves
- Seed Yield, including Number, Size, Weight, and/or Harvest Index
- Fruit Yield, including Number, Size, Weight, and/or Harvest Index
- Plant Development
- Time to Fruit Maturity
- Cell Wall Strengthening and Reinforcement
- Stress Tolerance
- Drought tolerance
- Flooding tolerance
- UV tolerance

To regulate any of the phenotype(s) above, activities of one or more of the phase responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in Anderson et al. (1997) *Plant Cell* 9: 1727-1743; Heintzen et al. (1997) *Proc. Natl. Acad. Sci. USA* 94: 8515-20; Schaffer et al. (1998) *Cell* 93:1219-1229; Somers et al. (1998) *Development* 125: 485-494; Somers et al. (1998) *Science* 282: 1488-1490; Wang and Tobin (1998) *Cell* 93: 1207-1217; Zhong et al. (1998) *Plant Cell* 10: 2005-2017; Sugano et al. (1998) *Proc. Natl. Acad. Sci. USA* 95: 11020-11025; Dowson-Day and Millar (1999) *Plant J* 17: 63-71; Green and Tobin (1999) *Proc. Natl. Acad. Sci. USA* 96: 4176-419; Staiger and Apel (1999) *Mol. Gen. Genet.* 261: 811-819; Strayer and Kay (1999) *Curr. Opin. Plant Biol.* 2:114-120; Strayer et. al. (2000)

*Science* 289:768-771; Kreps et al. (2000) *J Biol Rhythms* (2000) 15:208-217; Nelson et al. (2000) *Cell* 101:331-340; Somers et al. (2000) *Cell* 101:319-329.

b) USE OF PHASE RESPONSIVE GENES TO MODULATE  
BIOCHEMICAL ACTIVITIES

he activities of one or more of the phase responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations above and included in the table below:

Process	Biochemical Or Metabolic Activities And/Or Pathways	Citations including assays
Germination And Seedling Development	Cold, Light And Water Modulated Signal Transduction Pathways, Receptors, Kinases, PAS Domain Proteins	Bognar et al. (1999) <i>Proc. Natl. Acad. Sci. USA</i> 96:14652-14657; Sugano et al (1999) <i>Proc. Natl. Acad. Sci. USA</i> 96:12362-12366; Dowson-Day and Millar (1999) <i>Plant J</i> 17: 63-71; Somers et al. (2000) <i>Cell</i> 101:319-329; Zhong et al. (1998) <i>Plant Cell</i> 10: 2005-2017
Growth Transitions And Flowering	Cold And Light Modulated Signal Transduction Pathways, Receptors, Kinases, PAS Domain Proteins	Nelson et al. (2000) <i>Cell</i> 101:331-340; Fowler et al. (1999) <i>EMBO J.</i> 18:4679-4688
Tuber Formation	Cold And Light Modulated Signal Transduction Pathways	Yanovsky et al. (2000) <i>Plant J.</i> 23: 223-232

Process	Biochemical Or Metabolic Activities And/Or Pathways	Citations including assays
<u>METABOLISM</u>		
Lipid Metabolism	Membrane Lipid Synthesis Including Omega-3 Fatty Acid Desaturase, Lipases, Lipid Transfer Proteins	Bradley and Reddy (1997) <i>J. Bacteriol.</i> 179: 4407-4410; Martin, M et al. 1999 <i>Europe J. Biochem</i> 262: 283-290
Sugar Metabolism	Glycosylhydrolases, Glycosyltransferases, Amylases, Sucrose Synthase, CAB, Rubisco, Light Signal Transduction	Liu et al. (1996) <i>Plant Physiol.</i> 112:43-51; Millar and Kay (1996) <i>Proc Natl Acad Sci U S A</i> 93:15491-15496; Wang et al. (1997) <i>Plant Cell</i> 9:491-507; Shinohara et al (1999) <i>J. Biol. Chem.</i> 273: 446-452
Nitrogen Metabolism	Aminotransferases, Arginase, Proteases And Vegetative Storage Proteins, Aromatic Amino Acid Synthesis	Bradley and Reddy (1997) <i>J. Bacteriol.</i> 179: 4407-4410
Photorespiration	Mitochondrial, Chloroplast And Peroxisomal Photorespiratory Enzymes, Serine Hydroxymethyl Transferases, Catalase	Zhong and McClung (1996) <i>Mol. Gen. Genet.</i> 251:196-203; McClung (1997) <i>Free. Radic. Biol. Med.</i> 23:489-496; McClung et al. (2000) <i>Plant Physiol.</i> 123:381-392
Responses To Environmental Stress	Expression Of Genes Involved In Responses To Drought, Salt, UV	McClung (1997) <i>Free Radic Biol Med</i> 23:489-496; Shi et al. (2000) <i>Proc. Natl. Acad. Sci. USA</i> 97:6896-6901

Other biological activities that can be modulated by the phase responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Phase responsive genes are characteristically differentially transcribed in response to maturity of the cell, organ or tissue which depends on a timing mechanism, which is internal to an organism or cell. The Intensity Table reports the changes in transcript levels of various phase responsive genes in a plant.

The data from this experiment reveal a number of types of phase responsive genes and gene products. Profiles of some classes of phase responsive genes are shown in the table below with examples of which associated biological activities are modulated when the activities of one or more such genes vary in plants.

Transcript Levels	Type Of Genes	Physiological Consequences	Examples Of Biochemical Activity
Up Regulated Transcripts	<p>Responders To mutation that confers adult like phase</p> <p>Genes induced in adult-like phase</p>	<ul style="list-style-type: none"> <li>• Adult phase adoption</li> <li>• Metabolisms Affected By phase change</li> <li>• Synthesis Of Secondary Metabolites And/Or Proteins</li> <li>• Modulation Of Phase Response Transduction Pathways</li> <li>• Specific Gene Transcription Initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Metabolic Enzymes</li> <li>• Change In Cell Membrane Structure And Potential</li> <li>• Kinases And Phosphatases</li> <li>• Transcription Activators</li> <li>• Change In Chromatin Structure And/Or Localized DNA Topology</li> </ul>
Down-Regulated Transcripts	<p>Responders To mutation that confers adult phase</p> <p>Genes repressed in adult-like phase</p> <p>Genes With Discontinued</p>	<ul style="list-style-type: none"> <li>• Negative Regulation of adult phase pathways</li> <li>• Changes In Pathways And Processes Operating In Cells</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Change In Protein Structure By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>• Change In Chromatin Structure And/Or</li> </ul>

Transcript Levels	Type Of Genes	Physiological Consequences	Examples Of Biochemical Activity
	Expression Or Unstable mRNA in adult-like phase	<ul style="list-style-type: none"> <li>Changes In Metabolic pathways other than phase specific pathways</li> </ul>	<ul style="list-style-type: none"> <li>DNA Topology</li> <li>Stability Factors For Protein Synthesis And Degradation</li> <li>Metabolic Enzymes</li> </ul>

#### Use of Promoters of Phase Responsive Genes

Promoters of phase responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the phase responsive genes where the desired sequence is operably linked to a promoter of a phase responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.



## II. HORMONE RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

### II.A. ABSCISSIC ACID RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Plant hormones are naturally occurring substances, effective in very small amounts, which act as signals to stimulate or inhibit growth or regulate developmental processes in plants. Abscissic acid (ABA) is a ubiquitous hormone in vascular plants that has been detected in every major organ or living tissue from the root to the apical bud. The major physiological responses affected by ABA are dormancy, stress stomatal closure, water uptake, abscission and senescence. In contrast to auxins, cytokinins and gibberellins, which are principally growth promoters, ABA primarily acts as an inhibitor of growth and metabolic processes.

Changes in ABA concentration internally or in the surrounding environment in contact with a plant results in modulation of many genes and gene products. Examples of such ABA responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff, and MA\_clust tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA products changed in concentration in response to application of ABA to plants.

While ABA responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different ABA responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of an ABA responsive polynucleotide and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress and defence induced pathways, nutritional pathways and development. Here, in addition to polynucleotides having

similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways.

Such ABA responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in ABA concentration or in the absence of ABA fluctuations. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108560, 108561, 108513, 108597). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

ABA genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### ABA Genes Identified By Cluster Analyses Of Differential Expression

##### ABA Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of ABA genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108560, 108561, 108513, 108597 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### ABA Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of ABA genes. A group in the MA\_clust is considered a ABA pathway or network if the group comprises a cDNA ID that also appears in

Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

5

#### ABA Genes Identified By Amino Acid Sequence Similarity

ABA genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis ABA genes. Groups of ABA genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a ABA pathway or network is a group of proteins that also exhibits ABA functions/utilities.

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

Further, promoters of ABA responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by ABA or any of the following phenotypes or biological activities below.

#### II.A.1. USE OF ABSCISSIC ACID RESPONSIVE GENES TO MODULATE PHENOTYPES

ABA responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Development
  - Cell Growth
  - Promotion Of Leaf Cell Elongation
  - Fruit Development
  - Fruit Drop
  - Inhibition Of Parthenocarp And Ovary Growth
- Seed Development
  - Maturation Of Zygotic And Somatic Embryos
  - Embryo Development
  - Seed Development And Maturation
  - Acquisition Of Desiccation Tolerance

- Dormancy
  - Control Rate And Timing Of Germination
  - Prolongation Of Seed Storage And Viability
  - Inhibition Of Hydrolytic Enzyme Synthesis
- Growth
  - Roots
    - Inhibition Of Root Elongation Under Low Water Potential
  - Stems
  - Buds
    - Promotion Of Dormancy
    - Lateral/Axillary Bud Formation
  - Leaves
    - Inhibition Of ABA-Induced Growth And Elongation
- Biomass
  - Fresh And Dry Weight During Any Time In Plant Life, Such As Maturation;
  - Number, Size, And Weight Of
    - Flowers;
    - Seeds;
- Senescence
  - Abscission
  - Leaf Fall
  - Flower Longevity
- Differentiation
  - Plastid/Chloroplast Differentiation
  - Regulation Of Sterility
- Stress Responses
  - Mediation Of Response To Desiccation, Drought, Salt And Cold

To regulate any of the phenotype(s) above, activities of one or more of the ABA responsive genes or gene products can be modulated in an organism and tested by screening for

the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Koorneef and Karssen (1994, Seed dormancy and germination, In: Arabidopsis, Cold Spring harbor Lab. Press, pp 314-334), Cramer et al (1998, J. Exptl. Botany 49:191-198), and White and Rivin (2000, Plant Physiol 122: 1089-97). Phillips et al. (1997) EMBO J 16: 4489-96; Nambara et al (1995) Development 121: 629-636; Hays et al (1999) Plant Physiol. 119: 1065-72; Filonova et al (2000) J Exptl Botany 51: 249-64; White et al (2000) Plant Physiol. 122: 1081-88; and Visser et al. (1998) Plant Mol Biol 37: 131-40; Rohde et al. (2000) Plant Cell 12:35-52; and Cramer et al. (1998) J. experimental Botany. 49: 191-198.

#### II.A.2. USE OF ABSCISSIC ACID RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the ABA responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation And Development	<ul style="list-style-type: none"> <li>Farnesylation</li> <li>Nitrogen Metabolism</li> </ul>	<p>Pei Et Al (1998) Science 282: 287-290; Cutler Et Al. (1996) Science 273: 1239</p> <p>Goupil Et Al (1998) J Exptl Botany 49:1855-62</p>
Water Conservation And Resistance To Drought And Other Related Stresses	<ul style="list-style-type: none"> <li>Stomatal Development And Physiology</li> </ul>	<p>Allen Et Al. (1999) Plant Cell 11: 1785-1798</p> <p>Li Et Al. 2000 Science 287: 300-303</p> <p>Burnett Et Al 2000. J. Exptl Botany 51: 197-205</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	<ul style="list-style-type: none"> <li>Stress Response Pathways</li> <li>Inhibition Of Ethylene Production Under Low Water Potential</li> <li>Proline And Other Osmolite Synthesis And Degradation</li> </ul>	<p>Raschke (1987) In: Stomatal Function Zeiger Et Al. Eds., 253-279</p> <p>Bush And Pages (1998) Plant Mol. Biol. 37: 425-35</p> <p>Spollen Et Al (2000) Plant Physiol. 122:967-976</p> <p>Hare Et Al. (1998) Plant, Cell And Environment 21:535-553; Hare Et Al. (1999) J. Exptl. Botany 50:413-434</p>
	<ul style="list-style-type: none"> <li>Plasmalemma And Tonoplast Ion Channel Changes</li> <li>Ca<sup>2+</sup> Accumulation</li> <li>K<sup>+</sup> Efflux</li> <li>Activation Of Kinases And Phosphatases</li> </ul>	<p>Macrobbe (1998) Philos Trans R Soc Lond B Biol Sci 353: 1475-88; Li Et Al (2000) Science 287:300-303; Barkla Et Al. (1999) Plant Physiol. 120:811-819</p> <p>Lacombe Et Al. (2000) Plant Cell 12: 837-51; Wang Et Al. (1998) Plant Physiol 118:1421-1429; Shi Et Al. (1999) Plant Cell 11: 2393-2406</p> <p>Gaymard Et Al. (1998) Cell 94:647-655</p> <p>Jonak Et Al. (1996) Proc. Natl. Acad. Sci 93: 11274-79; Sheen (1998) Proc. Natl. Acad. Sci. 95: 975-80; Allen Et Al. (1999) Plant Cell 11: 1785-98</p>

Other biological activities that can be modulated by the ABA responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

5 ABA responsive genes are characteristically differentially transcribed in response to fluctuating ABA levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff reports the changes in transcript levels of various ABA responsive genes in entire seedlings at 1 and 6 hours after a plant was sprayed with a Hoagland's solution enriched with ABA as compared to seedlings sprayed with Hoagland's solution only.

10 The data from this time course can be used to identify a number of types of ABA responsive genes and gene products, including "early responders," and "delayed ABA responders", "early responder repressors" and "delayed repressors". Profiles of these different ABA responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts (Level At 1 Hr $\cong$ 6 Hr) or (Level At 1 Hr > 6 Hr)	Early Responders To ABA	ABA Perception ABA Uptake  Modulation Of ABA Response Transduction Pathways Specific Gene Transcription Initiation	Transcription Factors Transporters Change In Cell Membrane Structure Kinases And Phosphatases  Transcription Activators Change In Chromatin Structure And/Or Localized DNA Topology
Up Regulated Transcripts (Level At 1 Hr < 6 Hr)	Delayed Responders	Maintenance Of Response To ABA Maintenance Of Seed Dormancy, Stress Stomatal Closure, Water Uptake Control, Abscission And Senescence Control Pathways	Transcription Factors Specific Factors (Initiation And Elongation) For Protein Synthesis Maintenance Of Mrna Stability Maintenance Of Protein Stability Maintenance Of Protein-Protein Interaction
Down-Regulated Transcripts (Level At 1 Hr $\cong$ 6 Hr) or (Level At 6 Hr > 1 Hr)	Early Responder Repressors Of ABA State Of Metabolism	Negative Regulation Of ABA Pathways Released  Changes In Pathways	Transcription Factors Change In Protein Structure By Phosphorylation (Kinases) Or Dephosphorylation



TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Hr)	Genes With Discontinued Expression Or UnsTable mRNA In Presence Of ABA	And Processes Operating In Cells	(Phosphatases) Change In Chromatin Structure And/Or DNA Topology
Down-Regulated Transcripts (Level At 1 Hr > 6 Hr)	Delayed Repressors Of ABA State Of Metabolism  Genes With Discontinued Expression Or UnsTable mRNA In Presence Of ABA	Negative Regulation Of ABA Pathways Released  Maintenance Of Pathways Released From Repression  Changes In Pathways And Processes Operating In Cells	Transcription Factors Kinases And Phosphatases Stability Of Factors For Protein Synthesis And Degradation

#### Use of Promoters of ABA responsive Genes

Promoters of ABA responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the ABA responsive genes where the desired sequence is operably linked to a promoter of a ABA responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

Attorney Docket No. 2750-1481P

**Abstract** The purpose of this study was to determine the effect of a 12-week training program on the heart rate (HR) and blood pressure (BP) of sedentary, middle-aged men. The subjects were divided into two groups: a control group and an exercise group. The exercise group performed a 12-week training program consisting of aerobic and resistance exercises. The HR and BP were measured at baseline and at the end of the 12-week program. The results showed that the exercise group had a significant decrease in both HR and BP compared to the control group. The HR decreased from 72 to 68 beats per minute, and the BP decreased from 120/80 to 110/70 mmHg. The control group showed no significant changes in HR and BP. The findings suggest that a 12-week training program can effectively reduce HR and BP in sedentary, middle-aged men.

II.B. AUXIN RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Plant hormones are naturally occurring substances, effective in very small amounts that stimulate or inhibit growth or regulate developmental processes in plants. One of the plant hormones is indole-3-acetic acid (IAA), often referred to as Auxin.

Changes in Auxin concentration in the surrounding environment in contact with a plant or in a plant results in modulation of the activities of many genes and hence levels of gene products. Examples of such Auxin responsive genes and their products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. The genes were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to application of Auxin to plants.

Manipulation of one or more Auxin responsive gene activities are useful to modulate the biological activities and/or phenotypes listed below. Auxin response genes and gene products can act alone or in combination. Useful combinations include Auxin response genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of the levels of such proteins is especially useful for altering phenotypes and biochemical activities of plants. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108564, 108565, 108516, 108554, 108466, 107886, 107891, SMD 3743, and NAA (relating to SMD 3749, SMD 6338, SMD 6339)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

NAA genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

NAA Genes Identified By Cluster Analyses Of Differential Expression

NAA Genes Identified By Correlation To Genes That Are Differentially

Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of NAA genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108564, 108565, 108516, 108554, 108466, 107886, 107891, SMD 3743, and NAA (relating to SMD 3749, SMD 6338, SMD 6339) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

NAA Genes Identified By Correlation To Genes That Cause Physiological

Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of NAA genes. A group in the MA\_clust is considered a NAA pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

NAA Genes Identified By Amino Acid Sequence Similarity

NAA genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis NAA genes. Groups of NAA genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a NAA pathway or network is a group of proteins that also exhibits NAA functions/utilities.

Such Auxin responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in Auxin concentration or in the absence of Auxin fluctuations. Further, promoters of Auxin responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by auxin or any of the following phenotypes or biological activities below.

II.B.1. USE OF AUXIN RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE PHENOTYPES

Auxin Responsive Genes And Gene Products Are Useful To Or Modulate One Or  
More Of The Following Phenotypes:

- Growth
- Apical Dominance
- Vascular Growth
- Roots
- Inhibition Of Primary Root Elongation
- Increased Lateral Root Formation
- Stems
- Lateral Buds
- Lateral Branching
- Reduction Of Branching
- For High Density Growth Per Acre
- For Increased Wood Production
- Lateral Organ Initiation And/Or Positioning In Apical Meristem,
- Organ Formation, For Example, Fruit Number In Tomatoes
- Leaves
- Height/Stature, E.G., Taller Crops Or Increase Wood Production
- Regeneration And Differentiation Of Cultured Cells Or Plantlets
- Biomass
- Fresh And Dry Weight During Any Time In Plant Life, Such As Maturation;
- Number Of Flowers;
- Number Of Seeds;
- Number Of Branches;
- Number Of Leaves;
- Starch Content

- Seed Yield, Including Number, Size, Weight, Harvest Index, Starch Content
- Fruit Yield, Number, Size, Weight, Harvest Index, Starch Content
- Development
- Orienting Cell Growth,
- Establishment And Maintenance Of Plant Axis
- Apical Dominance
- Cell Plate Placement
- Polarised Growth, Initiation And/Or Development, Of Embryos Morphogenic Progression, E.G., From Early Radial To Late Axialized Torpedo Stages
- Differentiation Of Cells Into Morphologically Different Cell Layers
- Cotyledon Separation
- Fruit Development
- Abscission, Leading To Modulation Of Fruit Drop
- Parthenocarpy, Seedless Crops Resulting From Lack Of Seed Set
- Vascularization, E.G. Hypocotyl And Cotyledon Tissues
- Genetic Control Of Vascular Patterning And Influences Its Maturation;
- Specification Of The Sites Where Vascular Differentiation Will Occur;
- Determination Of The Direction And Extent Of Vascular Tissue Formation
- Maintenance Of The Continuity Of Vascular Development With Plant Growth
- Tropic Responses
- Gravitropic Responses, E.G. Affecting Roots And Shoots
- Modulation Of Phototropic Sensitivity, E.G. Increase Growth Under A Reduced Light Spectrum

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the Auxin responsive genes when the desired sequence is operably linked to a promoter of an Auxin responsive gene.

To modulate any of the phenotype(s) above, activities of one or more of the Auxin response genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be

screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance with Bechtold and Pelletier (1998). Methods Mol. Biol. 82: 259-266; Clough and Bent (1998). 16: 735-743; Krysan et al. (1999). Plant Cell 11:2283-2290.

## II.B.2. USE OF AUXIN RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS TO BIOCHEMICAL ACTIVITIES:

The activities of one or more of the Auxin responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below.

Such biological activities are documented and can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Growth and Differentiation	Protein Ubiquitination	Gray et al. (1999) Genes and Develop, 13:1678-1691 Bechtold and Pelletier (1998). Methods. Mol. Biol. 82:259-266
	Cell Wall loosening and Expansion	Catala et al. (2000). Plant Physiol. 122:527-534. Cosgrove, D. (1993). New Phytol. 124:1-23.

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Auxin/Cytokinin Ratio	Changing Auxin and/or cytokinin synthesis and/or turnover	Chen et al. (1988). Plant Physiol. 86:822-825 Tam et al. (2000). Plant Physiol. 123:589-595 Bartel and Fink. (1995). Science 268:1745-1748. Prinsen et al. (1995). Quantifying phytohormones in transformed plants. In: Methods in Molecular Biology. 44:245-262.
Auxin Transport	Channeling of polar Auxin Transport	Reed et al. (1998). Plant Physiol. 118:1369-1378. Estelle, M. (1998). Plant Cell 10:1775-1778
	Auxin Efflux Between Cells	Reed et al. (1998). Plant Physiol. 118:1369-1378. Marchant et al. (1999). EMBO J. 18:2066-2073.
	Auxin Influx In and Out of a Cell	Reed et al. (1998). Plant Physiol. 118:1369-1378. Marchant et al. (1999). EMBO J. 18:2066-2073.
	Electrogenic Proton Symport of Auxin	Young et al. (1999). Biochim Biophys Acta. 1415(2):306-22
Signal Transduction	K <sup>+</sup> Accumulation	Philippar et al. (1999). Proc. Natl. Acad. Sci. 96:12186-12191



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Permeability of Cell Membranes	Marchant et al. (1999). EMBO J. 18:2066-2073.
	Guanine-Nucleotide Exchange	Steinmann et al. (1999). Science 286:316-318. Peyroche et al. (1996). Nature 384:479-481.
	Protein Phosphorylation	Christensen et al. (2000). Cell 100:469-478. Hirt (2000). Proc. Natl. Acad Sci. 97:2405-2407.
	Interaction with Ethylene mode of action	Madlung et al. (1999). Plant Physiol. 120:897-906. Xu et al. (1998). Plant Physiol. 118:867-874.
Protein Turnover	Localization of Polypeptides with the basal End of Cells	Grebe et al. (2000). Plant Cell. 12:343-356

Other biological activities that can be modulated to by the Auxin responsive genes and their products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Domain section of the Reference Tables.

Auxin responsive genes are characteristically differentially transcribed in response to fluctuating Auxin levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff (s) report(s) the changes in transcript levels of various Auxin responsive genes in the aerial parts of a seedling at 1 and 6 hours after the seedling was sprayed with an solution enriched with Auxin as compared to aerial parts of a seedling sprayed with water.

The data from this time course can be used to identify a number of types of Auxin responsive genes and gene products, including "early responders," and "delayed

responders." Profiles of these different classes of Auxin responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

TRANSCRIPT LEVEL	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Upregulated transcripts (level at 1 hr $\leq$ 6 hours) (level at 1 hr > 6 hours)	Early responders to Auxin	<ul style="list-style-type: none"> <li>• Auxin perception</li> <li>• Auxin Uptake/transport</li> <li>• Modulation of Auxin response transduction pathways</li> <li>• Initiating transcription of specific gene(s)</li> <li>• Modification of cell walls</li> <li>• Modification of cell structures</li> <li>• Modification of metabolism</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Transporters; channeling of polar Auxin transport</li> <li>• Kinases and phosphatases; protein ubiquitination; guanine nucleotide exchange; changing Auxin and/or cytokinin synthesis and/or turnover; interaction with ethylene mode of action</li> <li>• Auxin metabolic pathways</li> <li>• Change in chromatin structure and/or DNA topology</li> <li>• Transcriptional activators</li> <li>• Change in activity of protein-protein interactions</li> <li>• Cell wall and cell growth promoting pathways</li> <li>• Change in activity of cytoskeletal proteins modulating cell structure</li> <li>• Metabolic enzymes</li> <li>• Coordination and control of central carbon and Auxin metabolism</li> </ul>

TRANSCRIPT LEVEL	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Upregulated transcripts (level at 1 hr <6 hr)	"Delayed" Responders	<ul style="list-style-type: none"> <li>• Completion and/or Maintenance of Auxin response</li> <li>• Initiating transcription of specific gene(s)</li> <li>• Modification of cell walls</li> <li>• Modification of cell structures</li> <li>• Modification of metabolism</li> </ul>	<p>Transcription factors</p> <p>Changes in membrane protein, membrane channel and/or transporter protein activity</p> <ul style="list-style-type: none"> <li>- Change in chromatin structure and/or DNA topology</li> <li>- Transcriptional activators</li> <li>- Change in activity of protein-protein interactions</li> <li>- Cell wall proteins</li> <li>- Change(s) in activity of cytoskeletal proteins modulating cell structure</li> <li>- Coordination and control of central carbon and Auxin metabolism</li> <li>-metabolic enzymes</li> </ul>

TRANSCRIPT LEVEL	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Downregulated transcripts (level at 1hour $\cong$ 6 hours) (level at 1hour > 6 hours)	Early repressor responders to Auxin  Genes for pathways diminished in presence of Auxin	<ul style="list-style-type: none"> <li>• Repression of Auxin induced proteins released</li> <li>• Reorientation of metabolism in certain cells</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Changes in activity of cytoskeletal proteins modulating cell structure</li> <li>• Changes in chromatin structure and/or DNA topology</li> <li>• Changes in protein structure and/or function by phosphorylation (kinases) and/or dephosphorylation (phosphatases)</li> <li>• Stability of factors for protein translation</li> <li>• Changes in cell membrane structure</li> <li>• Changes in chromatin and/or localized DNA topology</li> <li>• Changes in protein-protein interaction</li> <li>• Metabolic enzymes</li> </ul>

TRANSCRIPT LEVEL	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Down-regulated transcripts (level at 1 hour < 6 hours)	<p>"Delayed" repressor responders to Auxin</p> <p>Genes for pathways diminished in presence of Auxin</p>	<ul style="list-style-type: none"> <li>Maintenance of Auxin stimulated state(s) in certain cells</li> <li>Reorientation of metabolism in certain cells</li> </ul>	<ul style="list-style-type: none"> <li>Transcription factors</li> <li>Change in activity of cytoskeletal proteins modulating cell structure</li> <li>Changes in chromatin structure and/or DNA topology</li> <li>Changes in protein structure and/or function by phosphorylation (kinases) and/or dephosphorylation (phosphatases)</li> <li>Stability of factors for protein translation</li> <li>Changes in cell membrane structure</li> <li>Changes in chromatin and/or localized DNA topology</li> <li>Changes in protein-protein interaction</li> <li>Metabolic enzymes</li> </ul>

#### Use of Promoters of NAA responsive Genes

Promoters of NAA responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the NAA responsive genes where the desired sequence is operably linked to a promoter of a NAA responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

II.C. BRASSINOSTEROID RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS:

Plant hormones are naturally occurring substances, effective in very small amounts, which act as signals to stimulate or inhibit growth or regulate developmental processes in plants.

5 Brassinosteroids (BRs) are the most recently discovered, and least studied, class of plant hormones. The major physiological response affected by BRs is the longitudinal growth of young tissue via cell elongation and possibly cell division. Consequently, disruptions in BR metabolism, perception and activity frequently result in a dwarf phenotype. In addition, because BRs are derived from the sterol metabolic pathway, any perturbations to the sterol pathway can affect the BR pathway. In the same way, perturbations in the BR pathway can have effects on the later part of the sterol pathway and thus the sterol composition of membranes.

10 Changes in BR concentration in the surrounding environment or in contact with a plant result in modulation of many genes and gene products. Examples of such BR responsive genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant biomass and seed yield. These genes were discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA abundance changed in response to application of BRs to plants.

15 While BR responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different BR responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factors and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a BR responsive polynucleotide and/or gene product with another environmentally responsive polynucleotide is useful because of the interactions that exist between hormone regulated pathways, stress pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the

transcript levels of the experiment (see EXPT ID: 108580, 108581, 108557, 108478, 108479, 108480, 108481). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

BR genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### BR Genes Identified By Cluster Analyses Of Differential Expression

##### BR Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of BR genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108580, 108581, 108557, 108478, 108479, 108480, 108481 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### BR Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of BR genes. A group in the MA\_clust is considered a BR pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### BR Genes Identified By Amino Acid Sequence Similarity

BR genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis BR genes. Groups of BR



genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a BR pathway or network is a group of proteins that also exhibits BR functions/utilities.

Such BR responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in BR concentration or in the absence of BR fluctuations. Further, promoters of BR responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by BR or any of the following phenotypes or biological activities below.

#### II.C.1. USE OF BRASSINOSTEROID RESPONSIVE GENES TO MODULATE PHENOTYPES

Brassinosteroid responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
  - Promotes Cell Elongation
  - Elongation Accelerated At Low Temperatures For Increased Plant Growth In Marginal Lands
  - Acts In Concert With Other Hormones To Promote Cell Division
- Roots
  - Inhibitory To Root Growth
  - Expression In Roots Would Inhibit Bud Breaking Due To Higher Auxin:Cytokinin Ratio In Epicotyl
- Stems
  - Inhibits Radial Growth While Causing Stem Elongation
  - In Low Concentrations, Promotes Radial Expansion
  - Increases Biomass
- Height
- Seeds
  - Promotes Cell Expansion In Embryo And Thus Enhances Germination

- Leaves
- Increase Biomass
- Flowers
- Increase Reproduction
- Biomass
- Fresh And Dry Weight During Any Time In Plant Life, Such As Maturation;
- Number Of Flowers;
- Number Of Seeds;
- Number Of Branches;
- Number Of Leaves;
- Starch Content
- Seed Yield, Including Number, Size, Weight, Harvest Index, Starch Content
- Fruit Yield, Number, Size, Weight, Harvest Index, Starch Content
- Development
- Morphogenesis
- Control Of Organ Size And Shape
- Development Of New Ornamentals
- Control Of Leaf Size And Shape
- Promotes Leaf Unrolling And Enlargement
- For Development Of New Leafy Ornamentals
- Seed Development
- Inhibition Of De-Etiolation
- Dormancy
- Accelerated Germination At Low Temperatures
- Root
- Gravitropism
- Senescence
- Promoted In Light Grown Plants
- Inhibiting Synthesis Or Perception Could Extend Life Span Of Desired  
Tissues/Organs
- Differentiation

- Vascularization
- Promotes Xylem Differentiation
- Increases Xylem Fiber Length
- Resistance Responses
- Increases Resistance To Pathogens
- Tropic Responses

#### Gravitropic Responses Affecting Roots

Further, any desired sequence can be transcribed in similar temporal, tissue, or enviromentally specific patterns as the BR responsive genes when the desired sequence is operably linked to a promoter of a BR responsive gene.

To improve any of the desired phenotype(s) above, activities of one or more of the BR response genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266, and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50, visually inspected for the desired phenotype and metabolically and/or functionally assayed according to Choe et al. (1999, Plant Cell 11:207-21 and Plant Physiol 119: 897-907), Yamamoto et al. (1997, Plant Cell Physiol 38:980-3), Asami and Yshida (1999, Trends in Plant Sciences, 4:348-353) and Azpiroz et al. (1998, Plant Cell 10:219-230)

#### II.C.2. USE OF BRASSINOSTEROID RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the BR responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological

activities are documented and can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
BR Transport	<ul style="list-style-type: none"> <li>BR Efflux Between Cells</li> </ul>	B.Schulz and K. Feldmann, unpub. results
	<ul style="list-style-type: none"> <li>BR Influx In And Out Of A Cell</li> </ul>	B.Schulz and K. Feldmann, unpub. results
Signal Transduction	<ul style="list-style-type: none"> <li>Permeability Of Cell Membranes</li> </ul>	
	<ul style="list-style-type: none"> <li>Protein Phosphorylation</li> </ul>	
Metabolism	<ul style="list-style-type: none"> <li>Major Growth Coordinating Pathways</li> </ul>	

Other biological activities that can be modulated by the BR responsive genes and gene products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Domain section of the Reference Tables.

BR responsive genes are differentially transcribed in response to fluctuating BR levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various BR responsive genes in the aerial parts of a seedling at 1 and 6 hours after a plant was sprayed with a solution enriched with BR as compared to seedlings sprayed with water. The data from this time course can be used to identify a number of types of BR responsive genes and gene products, including "early responders," "delayed responders." Profiles of these different categories of BR responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts (Level At 1 Hr $\approx$ 6 Hr) (Level At 1 Hr $>$ 6 Hr)	<ul style="list-style-type: none"> <li>Early Responders To BR</li> </ul>	<ul style="list-style-type: none"> <li>BR Perception</li> <li>BR Transport</li> <li>BR Biosynthesis Feedback</li> <li>Modulation Of BR Response Transduction Pathways</li> <li>Specific Gene Transcription Initiation</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>Receptors</li> <li>Transporters</li> <li>Change In Cell Membrane Structure</li> <li>Feedback Regulated Biosynthetic Genes</li> <li>Kinases And Phosphatases</li> <li>2<sup>nd</sup> Messengers, Eg., Calmodulin</li> <li>Transcription Activators</li> <li>Change In Chromatin Structure And/Or Localized DNA Topology</li> </ul>
Up Regulated Transcripts (Level At 1 Hr $<$ 6 Hr)	<ul style="list-style-type: none"> <li>Delayed Responders</li> </ul>	<ul style="list-style-type: none"> <li>Maintenance Of Response To Br</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>BR Biosynthetic Genes</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		<ul style="list-style-type: none"> <li>Cell And Organ Elongation</li> <li>Gravitropism</li> </ul>	<ul style="list-style-type: none"> <li>Specific Factors (Initiation And Elongation) For Protein Synthesis</li> <li>Maintenance Of Mrna Stability</li> <li>Maintenance Of Protein Stability</li> <li>Maintenance Of Protein-Protein Interaction</li> <li>Cell Wall Elongation</li> </ul>
Down-Regulated Transcripts (Level At 1 Hr $\approx$ 6 Hr) (Level At 6 Hr $>$ 1 Hr)	Early Responder Repressors Of BR State Of Metabolism  Genes With Discontinued Expression Or Unstable Mrna In Presence Of BR	<ul style="list-style-type: none"> <li>Negative Regulation Of BR Pathways Released</li> <li>Changes In Pathways And Processes Operating In Cells</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>Change In Protein Structure By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>Change In Chromatin Structure And/Or DNA Topology</li> </ul>
Down-Regulated	<ul style="list-style-type: none"> <li>Delayed</li> </ul>	<ul style="list-style-type: none"> <li>Negative</li> </ul>	<ul style="list-style-type: none"> <li>Transcription</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Transcripts (Level At 1 Hr > 6 Hr)	Repressors Of BR State Of Metabolism  • Genes With Discontinued Expression Or Unstable Mra In Presence Of BR	Regulation Of BR Pathways Released  • Maintenance Of Pathways Released From Repression  • Changes In Pathways And Processes Operating In Cells	Factors • Kinases And Phosphatases • Stability Of Factors For Protein Synthesis And Degradation

#### Use of Promoters of BR responsive Genes

Promoters of BR responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the BR responsive genes where the desired sequence is operably linked to a promoter of a BR responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

II.D. CYTOKININ RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Plant hormones are naturally occurring substances, effective in very small amounts, which act as signals to stimulate or inhibit growth or regulate developmental processes in plants. Cytokinins (BA) are a group of hormones that are best known for their stimulatory effect on cell division, although they also participate in many other processes and pathways. All naturally occurring BAs are aminopurine derivatives, while nearly all synthetic compounds with BA activity are 6-substituted aminopurine derivatives. One of the most common synthetic BAs used in agriculture is benzylaminopurine (BAP).

Changes in BA concentration in the surrounding environment or in contact with a plant results in modulation of many genes and gene products. Examples of such BA responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to application of BA to plants.

While cytokinin responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different BA responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a BA responsive polynucleotide and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108566, 108567, 108517).



For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

BA genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### BA Genes Identified By Cluster Analyses Of Differential Expression

##### BA Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of BA genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108566, 108567, 108517 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### BA Genes Identified By Correlation To Genes That Cause Physiological

##### Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of BA genes. A group in the MA\_clust is considered a BA pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### BA Genes Identified By Amino Acid Sequence Similarity

BA genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis BA genes. Groups of BA genes are identified in the Protein Group table. In this table, any protein group that comprises a

peptide ID that corresponds to a cDNA ID member of a BA pathway or network is a group of proteins that also exhibits BA functions/utilities.

Such BA responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in BA concentration or in the absence of BA fluctuations.

Further, promoters of BA responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by BA or any of the following phenotypes or biological activities below.

#### II.D.1. USE OF BA-RESPONSIVE GENES TO MODULATE PHENOTYPES

BA responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
- Roots
  - Inhibition Of Elongation Of Root
- Stems
  - Inhibition Of Elongation Of Hypocotyl
- Lateral Buds
  - Promotion Of Outgrowth
    - For Rapid Production Of Multiple Shoots As A Source For Grafting
- Leaves
  - Development
    - Cell Growth
      - Expansion Of Cotyledon
      - Promotes Cell Enlargement
        - For Increased Yield From Leaf Crops
      - Chloroplast Development

- Delayed Degradation Of Chloroplasts
- For Increased Photosynthesis And Crop Yield
- Cell Division
  - For Increased Micropropagation
- Senescence
  - Delays
    - For Delayed Conversion From Photosynthesis To Salvage Programs In Leaves
    - For Increased Crop Yield
- Differentiation
  - Regulation Of Morphogenesis
    - For Manipulating Callus Growth And Shoot/Root Formation In Culture
- Maintenance Of Shoot Meristem
  - Increased Usable Wood Production
  - Reduced Tiller Number
    - For Denser Crop Planting Regimes
- Nutrient Metabolism
  - Effects On Seed Size
  - Effects On Rate Of Seed Set
    - For Increased Seed Yield
- Induction Of Ethylene Biosynthesis
  - Control Of Fruit Ripening
- Parthenocarp
  - Control Of Sexual Reproduction
  - Production Of Seedless Fruits

To regulate any of the phenotype(s) above, activities of one or more of the BA responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998,

Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or molecularly or metabolically or functionally assayed according to Lohman et al (1994, Physil. Plant 92:322-328), Woolhouse (1983, In Agricultural Research- Strategies of Plant reproduction, Meudt, ed., 201-236), Medford et al. (1989, Plant Cell 1: 403-13), Vogel et al. (1998, Genetics 149:417-27), Ehnes and Roitsch (1997, Plant J 1: 539-48), Rotino et al. (1997, Nat. Biotechnol. 15: 1398-1401).

## II.D.2. USE OF BA-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the BA responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Chloroplast Functioning	• Photosynthesis	Benkova et al (1999) Plant Physil 121: 245-252
Induction And Maintenance Of Cell Division	• Cell Cycle Phase Transition	Riou-Khamlichi et al. (1999) Science 283: 1541-44
Senescence	• Cell Death/Apoptosis	Lohman et al. (1994) Physiol Plant 92: 322-328
Signal Transduction	• Sensing Endogenous Stimuli To Trigger Growth And Shoot Formation	Kakimoto (1996) Science 274: 982-985

Other biological activities that can be modulated by the BA responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Domain section above.

5 BA responsive genes are characteristically differentially transcribed in response to fluctuating BA levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various BA responsive genes in the aerial parts of a seedling at 1 and 6 hours after a plant was sprayed with a Hoagland's solution enriched with BA as compared to seedlings sprayed with Hoagland's solution only.

10 The data from this time course can be used to identify a number of types of BA responsive genes and gene products, including "early responders," and "delayed responders." Profiles of these different BA responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENES	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts (Level At 1h $\approx$ 6h) Or (Higher At 1h Than 6h)	<ul style="list-style-type: none"> <li>Early Responders To BA</li> </ul>	<ul style="list-style-type: none"> <li>- BA Perception</li> <li>-BA Uptake</li> <li>-Modulation Of BA Response</li> <li>Transduction Pathways</li> <li>-Specific Gene Transcription Initiation</li> <li>-Initiate And Coordinate Cell Division</li> <li>-Regulation Of Pathways To</li> </ul>	<ul style="list-style-type: none"> <li>-Transcription Factors</li> <li>-Transporters</li> <li>-Kinase, Receptor-Like Protein Kinase</li> <li>-Ovule-Specific Homeotic Protein, Secretory Pathway</li> <li>-Cell Division Control Protein, Cyclins, Cyclin-Dependent Protein Kinase (Cdk), Cell Cycle Phosphatases, Mitosis-Specific Chromosome Segregation Protein, Mitotic Phosphoprotein, Dna Replication Proteins, Helicase Telomerase, Centromere Protein, tRNA Synthase</li> <li>-Senescence-Associated Protein, Bifunctional Nuclease, Aba Pathway Genes, Ethylene Pathway Genes, Proteases, Nucleases, Pcd Genes</li> </ul>

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENES	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
		Senescence	-Calvin Cycle, Chlorophyll A/B Binding Protein (Cab), Transketolase, Lipoxygenase, Chloroplast Rna Processing Protein, Chloroplast Envelope Membrane Protein.
		-Modulation Of Chloroplast Gene Expression And Photosynthesis	-Glutamate Synthase, Gogat, Asparagine Synthase, Catalase, Peroxidase
		-Modulation Of Photorespiration And Primary Nitrogen Assimilation In Leaves	-Heat Shock Proteins, Gst -Fatty Acid Elongase- Like Protein, Very-Long- Chain Fatty Acid Condensing Enzyme, Coa Synthase
		-Stress Response	-Vicilin Storage Protein -Homeobox Domain Proteins
		-Wax Biosynthesis	-Mutase, Phosphoglycerate Mutase -Pectate Lyase, Ethylene Pathway Genes

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENES	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
		-Nutrient Metabolism -Embryogenesis  -Glycolysis, Gluconeogenesis  -Ripening	
Upregulated Transcripts (Higher At 6h Than 1h)	BA Late Responders	-BA Responsive Pathways  -Cell Wall Extension  -Organogenesis  -Modulate Activation Of Disease Defense Genes  -Modulate Responses To External Stimuli  -Osmotic Stress Tolerance	-Transfactors, Kinases, Phosphatases, LRR's, Dna Remodelling Proteins, Cu-Binding Proteins  -Expansins, Extensins, Proline Rich Proteins  -AP2 Domain Containing Proteins -Transfactors Interacting With Resistant Genes -Glycin-Rich Proteins, Wall-Associated Receptor Kinase (Wak) -Proline Oxidase
Down-Regulated	Repressors Of BA	-Regulation Of	-Transfactors (Such As



Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the BA responsive genes when the desired sequence is operably linked to a promoter of a BA responsive gene.

II.E. GIBBERELIC ACID RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Plant hormones are naturally occurring substances, effective in very small amounts, which  
5 act as signals to stimulate or inhibit growth or regulate developmental processes in plants.

Gibberellic acid (GA) is a hormone in vascular plants that is synthesized in proplastids (giving  
rise to chloroplasts or leucoplasts) and vascular tissues. The major physiological responses  
affected by GA are seed germination, stem elongation, flower induction, anther development and  
seed and pericarp growth. GA is similar to auxins, cytokinins and gibberellins, in that they are  
principally growth promoters.

Changes in GA concentration in the surrounding environment or in contact with a plant  
result in modulation of many genes and gene products. Examples of such GA responsive genes  
and gene products are shown in the Reference and Sequence Tables. These genes and/or  
products are responsible for effects on traits such as plant vigor and biomass and seed yield.  
15 They were discovered and characterized from a much larger set of genes by experiments  
designed to find genes whose mRNA products changed in concentration in response to  
application of nitrogen to plants.

While GA responsive polynucleotides and gene products can act alone, combinations of  
these polynucleotides also affect growth and development. Useful combinations include  
20 different GA responsive polynucleotides and/or gene products that have similar transcription  
profiles or similar biological activities, and members of the same or similar biochemical  
pathways. Whole pathways and/or segments of pathways are controlled by transcription factors  
and proteins that affect the activity of signal transduction pathways. Therefore, manipulation of  
such protein levels is especially useful for altering phenotypes and biochemical activities of  
25 plants. In addition, the combination of a GA responsive polynucleotide and/or gene product with  
another environmentally responsive polynucleotide is also useful because of the interactions that  
exist between hormone regulated pathways, stress pathways, nutritional pathways and  
development. Here, in addition to polynucleotides having similar transcription profiles and/or  
biological activities, useful combinations include polynucleotides that may have different  
30 transcription profiles but which participate in common overlapping pathways. The MA\_diff  
and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript

levels of the experiment (see EXPT ID: 108562, 108563, 108519, 108520, 108521, 108484, 108485, 108486). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

GA genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### GA Genes Identified By Cluster Analyses Of Differential Expression

##### GA Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of GA genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108562, 108563, 108519, 108520, 108521, 108484, 108485, 108486 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### GA Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of GA genes. A group in the MA\_clust is considered a GA pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### GA Genes Identified By Amino Acid Sequence Similarity

GA genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis GA genes. Groups of GA

genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a GA pathway or network is a group of proteins that also exhibits GA functions/utilities.

Such GA responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in GA concentration or in the absence of GA fluctuations. Further, promoters of GA responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by GA or any of the following phenotypes or biological activities below.

#### II.E.1. USE OF GA RESPONSIVE GENES TO MODULATE PHENOTYPES:

GA responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
- Promotes Root Growth
- Promotes Cell Division
- Promotes Stem Elongation
- Secondary (Woody) Growth
- Promotes Growth In Leaves
- Biomass
- Increase In Stem And Leaf Mass
- Increase In Xylem Fiber Length And Biomass Production
- Development
- Cell Growth
- Fruit Development
- Seed Development
- Dormancy, Breaks Dormancy In Seeds And Buds
- Promotes Trichome Formation
- Decrease Senescence
- Regulation Of Fertility
- Stress Responses

- Flowering Time

Further, any desired sequence can be transcribed in similar temporal, tissue, or enviromentally specific patterns as the GA responsive genes when the desired sequence is operably linked to a promoter of a GA responsive gene.

To regulate any of the phenotype(s) above, activities of one or more of the GA response genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Hedden and Proebsting (1999, Plant Physiol. 119:365-370), Hedden and Phillips (1999, Current Opinion in Plant Biotech. 11:130-137), Perazza et al (1998, Plant Physiol. 117:375-383), Kende and Zeevaart (1997, Plant Cell 9:1197-1210) and van der Knaap et al. (2000, Plant Physiol. 122:695-704).

#### II.E.2. USE OF GA-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES:

The activities of one or more of the GA responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Growth and Differentiation	Biosynthesis of Gas	Hedden and Proebsting (1999, Plant Physiol. 119:365-370)
	Cell wall loosening and cell expansion	Cosgrove (1993, New Phytol. 124:1-23)

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	GA deactivation Major growth promoting metabolic pathways	Hedden and Proebsting (1999, Plant Physiol. 119:365-370)
Perception and Signal Transduction	Receptors	Koornneef and van der Veen (1980, TAG 58:257- 263)
	Synthesis of transcriptional regulators Calcium and Calmodulin	Bethke and Jones (1998, Curr. Opin. Plant Biol. 1:440-446)

Other biological activities that can be modulated by the GA responsive genes and gene products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Protein Domain table.

GA responsive genes are characteristically differentially transcribed in response to fluctuating GA levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various GA responsive genes in entire seedlings at 1 and 6 hours after a plant was sprayed with a Hoagland's solution enriched with GA as compared to seedlings sprayed with Hoagland's solution only.

The data from this time course can be used to identify a number of types of GA responsive genes and gene products, including "early responders," and "delayed responders." Profiles of some GA responsive genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
----------------------	---------------	-------------------------------	--

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts (level at 1 hr $\approx$ 6 hr) (level at 1 hr > 6 hr)	Early responders to GA  Genes induced by GA	GA perception GA transport Modulation of GA response transduction pathways Specific gene transcription initiation Growth stimulating pathway induction	Transcription factors Transporters Change in cell membrane structure Kinases and phosphatases  Transcription activators Change in chromatin structure and/or localized DNA topology Cell wall proteins Metabolic Enzymes
Up regulated transcripts (level at 1 hr < 6 hr)	Maintenance of GA response  "Delayed" responders	Maintenance of response to GA Induction of GA metabolic pathways	Transcription factors Specific factors (initiation and elongation) for protein synthesis Maintenance of mRNA stability Metabolic enzymes
Down-regulated transcripts (level at 1 hr $\approx$ 6 hr) (level at 6 hr > 1 hr)	Early repressor responders to GA  Genes repressed by GA	Negative regulation of GA pathways released  Reduced activity of repressed pathways	Transcription factors Calmodulin Change in protein structure by phosphorylation (kinases) or dephosphorylation

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	Genes whose activities are diminished or mRNAs are unstable in the presence of GA		(phosphatases) Change in chromatin structure and/or DNA topology
Down-regulated transcripts (level at 1 hr > 6 hr)	Delayed responders  Genes repressed by GA  Genes whose activities are diminished or mRNAs are unstable in the presence of GA	Maintenance or GA repressed pathways	Transcription factors Kinases and phosphatases Stability factors for protein translation Metabolic enzymes

#### Use of Promoters of GA responsive Genes

Promoters of GA responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the GA responsive genes where the desired sequence is operably linked to a promoter of a GA responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.



III. METABOLISM AFFECTING GENES, GENE COMPONENTS AND PRODUCTS

III.A. NITROGEN RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Nitrogen is often the rate-limiting element in plant growth, and all field crops have a  
5 fundamental dependence on exogenous nitrogen sources. Nitrogenous fertilizer which is usually  
supplied as ammonium nitrate, potassium nitrate, or urea, typically accounts for 40% of the costs  
associated with crops, such as corn and wheat in intensive agriculture. Increased efficiency of  
nitrogen use by plants should enable the production of higher yields with existing fertilizer inputs  
and/or enable existing yields of crops to be obtained with lower fertilizer input, or better yields on  
10 soils of poorer quality. Also, higher amounts of proteins in the crops could also be produced more  
cost-effectively.

Changes in nitrogen concentration in the surrounding environment or in contact with a plant  
results in modulation of the activities of many genes and hence levels of gene products. Examples  
of such "nitrogen responsive" genes and gene products with these properties are shown in the  
Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff  
MA\_diff and/or AFLP\_diff and/or AFLP\_diff, MA\_clust, Knock-in and Knock-out tables. These  
genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They  
were discovered and characterized from a much larger set by experiments designed to find genes  
whose mRNA products changed in response to changing levels of available nitrogen to plants.

20 Manipulation of one or more "nitrogen responsive" gene activities are useful to modulate  
the biological activities and/or phenotypes listed below. "Nitrogen responsive" genes and gene  
products can act alone or in combination. Useful combinations include nitrogen responsive  
genes and/or gene products with similar transcription profiles, similar biological activities, or  
members of the same or functionally related biochemical pathways. Whole pathways or  
25 segments of pathways are controlled by transcription factor proteins and proteins controlling the  
activity of signal transduction pathways. Therefore, manipulation of the levels of such proteins is  
especially useful for altering phenotypes and biochemical activities of plants. The MA\_diff  
and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript  
levels of the experiment (see EXPT ID: 108592, 108593, 108588, 108589, 108590, 108591,  
30 108532, 108548, 108549, 108550, 108551, 108454, 108455, 108487, 108488, 108489, and  
Nitrogen (relating to SMD 3787, SMD 3789)). For transcripts that had higher levels in the

samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Nitrogen genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Nitrogen Genes Identified By Cluster Analyses Of Differential Expression

##### Nitrogen Genes Identified By Correlation To Genes That Are Differentially

##### Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Nitrogen genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108592, 108593, 108588, 108589, 108590, 108591, 108532, 108548, 108549, 108550, 108551, 108454, 108455, 108487, 108488, 108489, and Nitrogen (relating to SMD 3787, SMD 3789) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Nitrogen Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Nitrogen genes. A group in the MA\_clust is considered a Nitrogen pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Nitrogen Genes Identified By Amino Acid Sequence Similarity

Nitrogen genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Nitrogen genes. Groups of Nitrogen genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Nitrogen pathway or network is a group of proteins that also exhibits Nitrogen functions/utilities.

Such “nitrogen responsive” genes and gene products can function either to either increase or dampen the phenotypes and activities below, either in response to changes in nitrogen concentration or in the absence of nitrogen fluctuations.

Further, promoters of nitrogen responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by nitrogen or any of the following phenotypes or biological activities below.

### III.A.1.USE OF NITROGEN-RESPONSIVE GENES TO MODULATE PHENOTYPES

“Nitrogen responsive” genes and gene products can be used to alter or modulate one or more of the following phenotypes:

- Plant Development
- Initiation of the Reproduction Cycle from a Vegetative State
  - Flower Development Time
  - Time to Fruit Maturity; and
- Root Development and Initiation
  - Root Branching
  - Lateral Root, Initiation and/or Development
  - Nodule formation and nitrogen Assimilation from any Nitrogen-Fixing Symbions.
- Growth Rate
- Whole Plant, including Height, Flowering Time, etc.
- Organs
  - Flowers
  - Fruits

- Stems
- Leaves
- Roots
- Lateral Roots
- 5      - Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation;
- Number, Size, and Weight of
  - Flowers;
  - Seeds;
  - 10      - Branches;
  - Leaves;
  - Total Plant Nitrogen Content
  - Amino Acid/Protein Content of Whole Plant or Parts
  - Seed Yield
  - 15      - Number, Size, Weight, Harvest Index
  - Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, and carbohydrate
  - Fruit Yield
  - Number, Size, Weight, Harvest Index
  - 20      -      Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, carbohydrate, Water

To regulate any of the phenotype(s) above, activities of one or more of the nitrogen responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant  
25      utilizing the procedures described herein and the phenotypes can be screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance to Zhang (1999) Proc. Natl. Acad. Sci. 96(11): 6529-34; or Zhang and Forde (1998) Science 279(5349):407-9; Scheible, W., Lauerer, M., Schultze, E.-D., Caboche, M., and Sitt, M. (1997). Plant J. 11, 671-691; Chevalier C, Bourgeois E, Just D, Raymond P. Plant J. 1996 Jan;9(1):1-  
30      11.

### III.A.2. USE OF NITROGEN-RESPONSIVE GENES TO MODULATE

#### BIOCHEMICAL ACTIVITIES

The activities of one or more of the nitrogen responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations included in the Table below:

Process	Biochemical Or Metabolic Activities And/Or Pathways	Citations including assays
Nitrate And Ammonium Uptake and Assimilation	NO <sub>3</sub> <sup>-</sup> Influx And Efflux	Lejay et al. (1999) Plant J. 18(5): 509-519
	Nitrate Channels	Liu et al. (1999) Plant Cell 11: 865-874; and Wang et al. (1998) Proc. Natl. Acad. Sci. USA 95: 15134-15139
	Changes In Membrane-Potential	Meharg et al. (1995) J. Membr. Biol. 145: 49-66; and Wang et al. (1998), supra
Amino Acid Synthesis	Glutamine Synthesis And Then Biosynthesis Of Other Amino Acids	Coruzzi et al. U.S. Pat. No. 5,955,651; and Oliveira et al. (1999) Plant. Phys. 121: 301-309
	Asparagine Synthesis And Then Biosynthesis Of Other Amino Acids	Lam et al. (1998) Plant J. 16(3): 345-353
Coordination Of Carbon And Nitrogen Metabolism	Light-Regulation Of Major Central Carbon And Nitrogen Metabolic Pathways To Coordinate Growth	Lam et al. (1998), supra; Lejay et al. (1999), supra; and Oliveira et al. (1999), supra

Process	Biochemical Or Metabolic Activities And/Or Pathways	Citations including assays
	Carbohydrate And Nitrogen Control Of Carbohydrate And Organic Nitrogen Accumulation Pathways	Lam et al. (1998) supra; Lejay et al. (1999) supra; and Oliveira et al. (1999) supra
Nitrogen Loading And Unloading	Nitrogen Transport From Source To Sinks	Walker et al. (1999) 210(1):9-18 Elsheikh et al. (1997) 51(2):137-44.
Nitrogen Storage	Accumulation Of Amino Acids And/Or Storage Proteins In Vacuoles	Johnson et al. (1990) Plant Cell 2(6):525-32.  Herman and Larkins (1999) Plant Cell. 11(4):601-14.
Ammonium Detoxification	Plastid Ammonium Storage/Glutamine Synthesis	Crawford (1995) Plant Cell 7(7):859-68. Zhang and Forde (1998) Science 279: 407-409.
Cell Growth	Division And/Or Elongation	Zhang and Forde (1998) Science 279: 407-409.  Coruzzi et al. U.S. Pat. No. 5,955,651

Other biological activities that can be modulated by the nitrogen responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Domain section above.

Nitrogen responsive genes are characteristically differentially transcribed in response to fluctuating nitrogen levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various nitrogen responsive genes in the aerial parts of a seedling at 2, 6, 9 and 12 hours after a plant was sprayed with a solution enriched with ammonium nitrate as compared to seedlings sprayed with water. The MA\_diff and/or AFLP\_diff MA\_diff and/or

AFLP\_diff and/or AFLP\_diff reports the changes in transcript levels of various nitrogen responsive genes in roots at 12 and 24 hours that were cut from seedlings transferred from a high to low potassium nitrate environment compared to control seedlings transferred to a high potassium nitrate environment.

- 5       The data from this time course reveal a number of types of nitrogen responsive genes and gene products, including “early responders,” and “delayed nitrogen responders”. Profiles of the individual categories of nitrogen responsive genes are shown in the Tables below together with examples of the kinds of associated biological activities that are modulated when the activities of one or more such genes vary in plants.

Low to High Ammonium Nitrate Experiment

Gene Expression Levels	Functional Category Of Gene	Physiological Consequences	Examples Of Gene Products
Upregulated Transcripts (Level At 2h $\geq$ 6, 9 Or 12h) Or (Level At 2h > 6, 9 Or 12h)	Early Responders To Nitrogen	<ul style="list-style-type: none"> <li>- Perception Of Nitrogen</li> <li>- Induced Nitrogen Uptake Into Cells</li> <li>- Induction Of Nitrogen Response Transduction Pathways</li> <li>- Initiation Of Specific Gene Transcription</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Transporters</li> <li>- Inhibitors Of Nitrogen Fixation</li> <li>- Components Of Pathways Released From Repression</li> <li>- Transaminases</li> <li>- Amino Acid Biosynthetic Enzymes</li> </ul>
Upregulated Transcripts (Level At 2h < 6, 9, Or 12h)	Delayed Nitrogen Responders	<ul style="list-style-type: none"> <li>- Maintenance Of High Nitrogen Metabolism And Growth</li> </ul>	<ul style="list-style-type: none"> <li>- Nitrogen Metabolic Pathway Enzymes</li> <li>- Transaminases</li> <li>- Amino Acid Biosynthetic Enzymes</li> <li>- Factors Induced In Coordination And Control Of Central Carbon And Nitrogen Metabolism</li> <li>- Cell Wall And Cell Growth- Promoting Pathway Enzymes</li> <li>- Storage Proteins</li> </ul>



Gene Expression Levels	Functional Category Of Gene	Physiological Consequences	Examples Of Gene Products
Down Regulated Transcripts (Level At 2h $\approx$ 6, 9 Or 12h) Or (Level At 6, 9 Or 12h > 2h)	- Early Responder Repressors Of Nitrogen Utilization Pathways  - Genes With Discontinued Expression Or Unstable Mrna Following Nitrogen Uptake	- Negative Regulation Of Nitrogen Utilization Pathways Released  - Pathways Of C And N Metabolism Required At Lower Levels Decline In Presence Of High Nitrogen	- Transcription Factors - Kinases And Phosphatases - Cytoskeletal Proteins Modulating Cell Structure - Chromatin Structure Regulatory Proteins - Metabolic Enzymes -Transporters - Proteins And Rna Turnover Systems
Level At 2 Hours > 6,9 Or 12 Hours	- Delayed Response Repressors Of Nitrogen Utilization Pathways - Genes With Discontinued Expression Or Unstable Mrna Following Nitrogen Uptake	- Negative Regulation Of Nitrogen Utilization Pathways Released  - Pathways Of C And N Metabolism Required At Lower Levels Decline In Presence Of High Nitrogen	- Transcription Factors - Kinases And Phosphatases - Cytoskeletal Proteins Modulating Cell Structure - Chromatin Structure Regulatory Proteins - Metabolic Enzymes - Transporters - Protein And Rna Turnover Systems

High to Low Potassium Nitrate Experiments

Gene Expression Levels	Functional Category Of Gene	Type Of Biological Activity	examples Of Biochemical Activities Of Gene Products
Upregulated Transcripts (Level At 12h ~ 24h) (Level At 12h>24h)	Early Responders To Low Nitrate	<ul style="list-style-type: none"> <li>- Perception Of Low Nitrate</li> <li>- Nitrogen Uptake Into Cells</li> <li>- Low Nitrogen Signal Transduction Response Pathways</li> <li>- Initiation Of Specific Gene Transcription</li> <li>- Initiation Of Nitrogen Fixation</li> </ul>	<p>Transcription Factors – Controlling Transcription</p> <p>Transporters – Facilitating Transport</p> <p>Cell Wall/Membrane structure Determining proteins</p> <p>Kinases And Phosphatases- regulating Signal transduction Pathways</p> <p>Cytoskeletal Proteins- Modulating Cell Structure</p> <p>Chromatin Structure And/Or Dna Topology Proteins</p> <p>Protein-Protein Interaction participants</p> <p>Metabolic Enzymes- nitrogen Turnover Enzymes and Pathway Components</p>

Gene Expression Levels	Functional Category Of Gene	Type Of Biological Activity	Examples Of Biochemical Activities Of Gene Products
Upregulated Transcripts (Level 12h<24h)	Delayed Low Nitrate Responders	- Maintenance Of Low Nitrogen Response Pathways (See the Table Above)	Transcription Factors – Controlling Transcription Transporters – Facilitating Transport Cell Wall/Membrane structure Determining proteins Kinases And Phosphatases- regulating Signal transduction Pathways Cytoskeletal Proteins- Modulating Cell Structure Chromatin Structure And/Or Dna Topology Proteins Protein-Protein Interaction participants Metabolic Enzymes- Nitrogen Turnover Enzymes and Pathway Components

Gene Expression Levels	Functional Category Of Gene	Type Of Biological Activity	examples Of Biochemical Activities Of Gene Products
Down-Regulated Transcripts (Level At 12h~24h) (Level At 12h>24h)	- Early Repressor Responders To Low Nitrate  - Genes Whose Expression Is Discontinued Or Mrna Is UnsTable In Presence Of Low Nitrate	-Negative Regulation Of Low Nitrogen-Mediated Pathways And/Or Responses Released - Pathways In C And N Metabolism Required At Lower Levels Decline In The Presence Of Low Nitrate	- Transcription Factors  Cell Wall/Membrane Structure Determining Proteins  Factors For Promoting Protein Translation  Kinases And Phosphatases  Cytoskeletal Proteins- odulating Cell Structure Protein And Rna Turnover systems

Gene Expression Levels	Functional Category Of Gene	Type Of Biological Activity	examples Of Biochemical Activities Of Gene Products
Down-Regulated Transcripts (Level At 12h<24h)	- Delayed Repressor Responders To Low Nitrate  - Genes Whose Expression Is Discontinued Or mRNA Is Unstable In Presence Of Low Nitrate	Negative Regulation Of Low Nitrogen-Mediated Pathways And/OR Responses Released Pathways In C And N Metabolism Required At Lower Levels Decline In The Presence Of Low Nitrate	Transcription Factors  Cell Wall/Membrane Structure Determining Proteins  Factors For Promoting Protein Translation  Kinases And Phosphatases  Cytoskeletal Proteins- modulating Cell Structure Protein And Rna Turnover systems  Chromatin Structure and/or Dna Topology proteins

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the nitrogen responsive genes when the desired sequence is operably linked to a promoter of a nitrogen responsive gene.

III.B. CIRCADIAN RHYTHM RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Often growth and yield are limited by the ability of a plant to tolerate stress conditions,  
including water loss. To combat such conditions, plant cells deploy a battery of responses that  
are controlled by an internal circadian clock, including the timed movement of cotyledons and  
leaves, timed movements in guard cells in stomata, and timed biochemical activities involved  
with sugar and nitrogen metabolism. These responses depend on the functioning of an internal  
circadian clock, that becomes entrained to the ambient light/dark cycle, and a series of  
downstream signaling events leading to transcription independent and transcription dependent  
stress responses.

A functioning circadian clock can anticipate dark/light transitions and prepare the  
physiology and biochemistry of a plant accordingly. For example, expression of a chlorophyll  
a/b binding protein (CAB) is elevated before daybreak, so that photosynthesis can operate  
maximally as soon as there is light to drive it. Similar considerations apply to light/dark  
transitions and to many areas of plant physiology such as sugar metabolism, nitrogen  
metabolism, water uptake and water loss, flowering and flower opening, epinasty, germination,  
perception of season, and senescence.

Manipulation of one or more clock gene activities are useful to modulate the biological  
processes and/or phenotypes listed below. Clock responsive genes and gene products can act  
alone or in combination. Useful combinations include clock responsive genes and/or gene  
products with similar transcription profiles, similar biological activities, or members of the same  
or functionally related biochemical pathways. Whole pathways or segments of pathways are  
controlled by transcription factor proteins and proteins controlling the activity of signal  
transduction pathways. Therefore, manipulation of such protein levels is especially useful for  
altering phenotypes and biochemical activities of plants. The MA\_diff and/or AFLP\_diff  
MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the  
experiment (see EXPT ID: Circadian (relating to SMD 2344, SMD 2359, SMD 2361, SMD  
2362, SMD 2363, SMD 2364, SMD 2365, SMD 2366, SMD 2367, SMD 2368, SMD 3242)).  
For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is

shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Circadian genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Circadian Genes Identified By Cluster Analyses Of Differential Expression

##### Circadian Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Circadian genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Circadian (relating to SMD 2344, SMD 2359, SMD 2361, SMD 2362, SMD 2363, SMD 2364, SMD 2365, SMD 2366, SMD 2367, SMD 2368, SMD 3242) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Circadian Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Circadian genes. A group in the MA\_clust is considered a Circadian pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Circadian Genes Identified By Amino Acid Sequence Similarity

Circadian genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Circadian genes.

Groups of Circadian genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Circadian pathway or network is a group of proteins that also exhibits Circadian functions/utilities.

Such clock responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in daylength or in response to changes in light quality. Further, promoters of circadian (clock) responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by circadian or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the circadian (clock) responsive genes when the desired sequence is operably linked to a promoter of a circadian (clock) responsive gene.

The expression of many genes is modulated by the clock. Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing some 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in response to the circadian rhythm clock at various times through the 24 hour cycle relative to the controls were identified, compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which represent clock responsive genes.

### III.B.1. USE OF CIRCADIAN RHYTHM (CLOCK) RESPONSIVE GENES TO MODULATE PHENOTYPES



Clock responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Timing Phenotypes
- Dormancy
- Germination
- Cotyledon opening
- First leaves
- Juvenile to adult transition
- Bolting
- Flowering
- Pollination
- Fertilization
- Seed development
- Seed set
- Fruit Drop
- Senescence
- Epinasty
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation
- Number of Flowers, Seeds, Branches, and/or Leaves
- Seed Yield, including Number, Size, Weight, and/or Harvest Index
- Fruit Yield, including Number, Size, Weight, and/or Harvest Index
- Plant Development
- Time to Fruit Maturity
- Cell Wall Strengthening and Reinforcement
- Stress Tolerance
- Drought tolerance
- Flooding tolerance
- UV tolerance

To regulate any of the phenotype(s) above, activities of one or more of the clock responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in

Anderson et al. (1997) Plant Cell 9: 1727-1743; Heintzen et al. (1997) Proc. Natl. Acad. Sci. USA 94: 8515-20; Schaffer et al. (1998) Cell 93:1219-1229; Somers et al. (1998) Development 125: 485-494; Somers et al. (1998) Science 282: 1488-1490; Wang and Tobin (1998) Cell 93: 1207-1217; Zhong et al. (1998) Plant Cell 10: 2005-2017; Sugano et al. (1998) Proc. Natl. Acad. Sci. USA 95: 11020-11025; Dowson-Day and Millar (1999) Plant J 17: 63-71; Green and Tobin (1999) Proc. Natl. Acad. Sci. USA 96: 4176-419; Staiger and Apel (1999) Mol. Gen. Genet. 261: 811-819; Strayer and Kay (1999) Curr. Opin. Plant Biol. 2:114-120; Strayer et. al. (2000) Science 289:768-771; Kreps et al. (2000) J Biol Rhythms (2000) 15:208-217; Nelson et al. (2000) Cell 101:331-340; Somers et al. (2000) Cell 101:319-329.

### III.B.2. USE OF ACTIVE CLOCK RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the clock responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations above and included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Germination and seedling development	Cold, light and water modulated signal transduction pathways, receptors, kinases, PAS domain	Bognar et al. (1999) Proc. Natl. Acad. Sci. USA 96:14652-14657; Sugano et al (1999) Proc. Natl. Acad. Sci. USA 96:12362-12366; Dowson-Day and Millar (1999) Plant J 17: 63-71; Somers et al. (2000) Cell 101:319-329; Zhong et al. (1998) Plant Cell 10: 2005- 2017
Growth transitions and flowering	Cold and light modulated signal transduction pathways, receptors, kinases, PAS domain	Nelson et al. (2000) Cell 101:331-340; Fowler et al. (1999) EMBO J. 18:4679- 4688
Tuber formation	Cold and light modulated signal transduction pathways	Yanovsky et al. (2000) Plant J. 23: 223-232
<u>METABOLISM</u>		
Lipid metabolism	Membrane lipid synthesis including omega-3 fatty acid desaturase, lipases, lipid transfer proteins	Bradley and Reddy (1997) J. Bacteriol. 179: 4407-4410; Martin, M et al. 1999 Europe J. Biochem 262: 283-290

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Sugar metabolism	Glycosylhydrolases, glycosyltransferases, amylases, sucrose synthase, CAB, Rubisco, light signal transduction	Liu et al. (1996) Plant Physiol. 112:43-51; Millar and Kay (1996) Proc Natl Acad Sci U S A 93:15491- 15496; Wang et al. (1997) Plant Cell 9:491-507; Shinohara et al (1999) J. Biol. Chem. 273: 446-452
Nitrogen metabolism	Aminotransferases, arginase, proteases and vegetative storage proteins, aromatic amino acid synthesis	Bradley and Reddy (1997) J. Bacteriol. 179: 4407-4410
Photorespiration	Mitochondrial, chloroplast and peroxisomal photorespiratory enzymes, serine hydroxymethyl transferases, catalase	Zhong and McClung (1996) Mol. Gen. Genet. 251:196- 203; McClung (1997) Free. Radic. Biol. Med. 23:489- 496; McClung et al. (2000) Plant Physiol. 123:381-392
Responses to Environmental Stress	Expression of genes involved in responses to drought, salt, UV	McClung (1997) Free Radic Biol Med 23:489-496; Shi et al. (2000) Proc. Natl. Acad. Sci. USA 97:6896-6901

Other biological activities that can be modulated by the clock responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Clock responsive genes are characteristically differentially transcribed in response to fluctuations in an entrained oscillator, which is internal to an organism and cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various clock responsive genes in a plant.

5 Profiles of clock responsive genes are shown in the table below with examples of which associated biological activities are modulated when the activities of one or more such genes vary in plants.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	<p>Responders to circadian rhythm</p> <p>Genes induced by rhythm</p>	<ul style="list-style-type: none"> <li>• Circadian rhythm perception</li> <li>• Metabolisms affected by Circadian rhythm</li> <li>• Synthesis of secondary metabolites and/or proteins</li> <li>• Modulation of clock response transduction pathways</li> <li>• Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Metabolic enzymes</li> <li>• Change in cell membrane structure and potential</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> <li>• Enzymes in lipid, sugar and nitrogen metabolism</li> <li>• Enzymes in photorespiration and photosynthesis</li> </ul>

Down-regulated transcripts	<p>Responders to circadian rhythm.</p> <p>Repressors of circadian "state" of metabolism</p> <p>Genes repressed by rhythm</p> <p>Genes with discontinued expression or unstable mRNA in presence of zinc</p>	<ul style="list-style-type: none"> <li>• Negative regulation of circadian pathways released</li> <li>• Changes in pathways and processes operating in cells</li> <li>• Changes in metabolism other than circadian pathways</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> <li>• Change in chromatin structure and/or DNA topology</li> <li>• Stability of factors for protein synthesis and degradation</li> <li>• Metabolic enzymes in light, sugar, lipid and nitrogen metabolism</li> </ul>
----------------------------	---	---	--

#### Use of Promoters of Clock responsive Genes

Promoters of Clock responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Clock responsive genes where the desired sequence is operably linked to a promoter of a Clock responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

III.C. BLUE LIGHT (PHOTOTROPISM) RESPONSIVE GENES, GENE  
COMPONENTS AND PRODUCTS

Phototropism is the orientation or growth of a cell, an organism or part of an organism in  
5 relation to a source of light. Plants can sense red (R), far-red (FR) and blue light in their  
environment and respond differently to particular ratios of these. For example, a low R:FR ratio  
enhances cell elongation and favors flowering over leaf production, but blue light regulated  
cryptochromes also appear to be involved in determining hypocotyl growth and flowering time.

Phototropism of *Arabidopsis thaliana* seedlings in response to a blue light source is  
10 initiated by nonphototropic hypocotyl 1 (NPH1), a blue light-activated serine-threonine protein  
kinase, but the downstream signaling events are not entirely known. Blue light treatment leads  
to changes in gene expression. These genes have been identified by comparing the levels of  
mRNAs of individual genes in dark-grown seedlings, compared with in dark grown seedlings  
treated with 1 hour of blue light. Auxin also affects blue light phototropism. The effect of auxin  
15 on gene expression stimulated by blue light has been explored by studying mRNA levels in a  
mutant of *Arabidopsis thaliana* *nph4-2*, grown in the dark and, treated with blue light for 1 hour  
compared with wild type seedlings treated similarly. This mutant is disrupted for auxin-related  
growth and auxin-induced gene transcription. Gene expression was studied using microarray  
technology.

20 Microarray technology allows monitoring of gene expression levels for thousands of  
genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to  
glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US  
*Arabidopsis* Functional Genomics Consortium (AFGC) has recently made public the results from  
such microarray experiments conducted with AFGC chips containing some 10,000 non-  
25 redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA  
of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases over the  
controls were identified, compared to the Ceres full-length cDNA and genomic sequence  
databanks, and the equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff  
30 and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the results of this analysis, indicating  
those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres

clones which represent blue light responsive genes and of those which are blue light responsive in the absence of nph4 gene activity. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Phototropism (relating to SMD 4188, SMD 6617, SMD 6619)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Blue Light genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Blue Light Genes Identified By Cluster Analyses Of Differential Expression

##### Blue Light Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Blue Light genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Phototropism (relating to SMD 4188, SMD 6617, SMD 6619) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Blue Light Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Blue Light genes. A group in the MA\_clust is considered a Blue Light pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.



Blue Light Genes Identified By Amino Acid Sequence Similarity

Blue Light genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Blue Light genes.

Groups of Blue Light genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Blue Light pathway or network is a group of proteins that also exhibits Blue Light functions/utilities.

III.C.1.1.USE OF BLUE LIGHT RESPONSIVE GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE PHENOTYPES

Changes in blue light in a plant's surrounding environment result in modulation of many genes and gene products. Examples of such blue light response genes and gene products are shown in the REFERENCE and SEQUENCE Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield.

While blue light responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different blue light responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a blue light responsive polynucleotides and/or gene product with other environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress and pathogen induced pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways.

III.C.2. USE OF BLUE LIGHT RESPONSIVE GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE PHENOTYPES

Blue light responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in blue light response concentration or in the absence of blue light responsive fluctuations. Further, promoters of blue light responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by blue light or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the blue light responsive genes when the desired sequence is operably linked to a promoter of a blue light responsive gene.

Blue light responsive genes and gene products can be used to alter or modulate one or more of the following phenotypes:

- Growth
  - Roots
    - Elongation
    - gravitropism
  - Stems
    - Elongation
- Development
  - Cell
    - Growth
    - Elongation
  - Flower
    - Flowering time
  - Seedling
    - Elongation
  - Plant Yield
  - Seed and fruit yield

To regulate any of the phenotype(s) above, activities of one or more of the blue light responsive genes or gene products can be modulated and the plants tested by screening for

the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Liscum and Briggs (1995, Plant Cell 7: 473-85), Vitha et al. (2000, Plant Physiol 122: 453-61), Stowe-Evance et al. (1998, Plant Physiol 118: 1265-75), Baum et al. (1999, PNAS USA 96: 13554-9), Huala et al. (1997) Science 278: 2120-2123), Kanegae et al. (2000, Plant Cell Physiol 41: 415-23), Khanna et al. (1999, Plant Mol Biol 39: 231-42), Sakai et al. (2000, Plant Cell 12: 225-36), Parks et al (1996, Plant Physiol 110: 155-62) and Janoudi et al. (1997, Plant Physiol 113: 975-79).

### III.C.3. USE OF BLUE LIGHT RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the blue light responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Growth and Development	Cell Elongation <ul style="list-style-type: none"> <li>Seedling</li> <li>Stem</li> </ul>	Liscum and Briggs (1995) Plant Cell 7: 473-85
	Root	Vitha et al. (2000) Plant Physiol 122: 453-61
Signalling	UV light Perception	Liscum and Briggs (1996) Plant Physiol 112: 291-96

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Far-red/Red light Perception	Parks et al. (1996) Plant Physiol 110: 155-62
	Phosphorylation of cellular and nuclear-localized proteins	Liscum and Briggs (1996) Plant Physiol 112: 291-96
	Activation and Synthesis of Transcription Factors	Sakae et al. (2000) Plant Cell 12: 225-36
	Ca <sup>2+</sup> levels	Baum et al. (1999) PNAS USA 96: 13554-9 Pu and Robinson (1998) J Cell Sci 111: 3197-3207
	Auxin Concentration	Estelle (1998) Plant Cell 10: 1775-8 Reed et al. (1998) Plant Physiol 118: 1369-78
	Inter-photoreceptors	Janoudi et al. (1997) Plant Physiol 113: 975-79

Other biological activities that can be modulated by blue light response genes and their products are listed in the REF Tables. Assays for detecting such biological activities are described in the Domain section of the REF Table.

5

The specific genes modulated by blue light, in wild type seedlings and in the mutant deficient in transmitting auxin effects are given in the Reference and Sequence Tables. The kinds of genes discovered and some of their associated effects are given in the Table below.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	Responders to no blue light in wild type or to blue light in mutant lacking auxin effects	<ul style="list-style-type: none"> <li>• Blue light perception</li> <li>• Metabolism affected by blue light</li> <li>• Synthesis of secondary metabolites and/or proteins</li> <li>• Modulation of blue light transduction pathways</li> <li>• Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic enzymes</li> <li>• Change in cell membrane structure and potential</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> </ul>
Down-regulated transcripts	<p>Responders to no blue light in wild type or to blue light in mutants lacking auxin effects</p> <p>Genes with discontinued expression or unsTable mRNA during response</p>	<ul style="list-style-type: none"> <li>• Blue light perception</li> <li>• Metabolism affected by blue light</li> <li>• Synthesis of secondary metabolites and/or proteins</li> <li>• Modulation of blue light transduction pathways</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> <li>• Change in chromatin structure and/or DNA topology</li> <li>• Stability factors for protein synthesis and</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		<ul style="list-style-type: none"> <li>• Specific gene transcription initiation</li> <li>• Changes in pathways and processes operating in cells</li> <li>• Changes in metabolic pathways other than phototropic blue light responsive pathways</li> </ul>	<ul style="list-style-type: none"> <li>• degradation</li> <li>• Metabolic enzymes</li> </ul>

#### Use of Promoters of Blue Light responsive Genes

Promoters of Blue Light responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Blue Light responsive genes where the desired sequence is operably linked to a promoter of a Blue Light responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

III.D. RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

There has been a recent and significant increase in the level of atmospheric carbon dioxide. This rise in level is projected to continue over the next 50 years. The effects of the increased level of carbon dioxide on vegetation are just now being examined, generally in large scale, whole plant (often trees) experiments. Some researchers have initiated physiological experiments in attempts to define the biochemical pathways that are either affected by and/or are activated to allow the plant to avert damage from the elevated carbon dioxide levels. A genomics approach to this issue, using a model plant system, allows identification of those pathways affected by and/or as having a role in averting damage due to the elevated carbon dioxide levels and affecting growth. Higher agronomic yields can be obtained for some crops grown in elevated CO<sub>2</sub>.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The U.S. Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing about 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in plants grown in higher CO<sub>2</sub> levels compared with plants grown at more normal CO<sub>2</sub> levels, were compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones were identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones cDNA sequences that change in response to CO<sub>2</sub>.

Examples of CO<sub>2</sub> responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. While CO<sub>2</sub> responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also

affect growth and development. Useful combinations include different CO<sub>2</sub> responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants.

Manipulation of one or more CO<sub>2</sub> responsive gene activities are useful to modulate the biological processes and/or phenotypes listed below. CO<sub>2</sub> responsive genes and gene products can act alone or in combination. Useful combinations include genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways.

CO<sub>2</sub> responsive genes and gene products can function to either increase or dampen the above phenotypes or activities. Further, promoters of CO<sub>2</sub> responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by CO<sub>2</sub> or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the CO<sub>2</sub> responsive genes when the desired sequence is operably linked to a promoter of a CO<sub>2</sub> responsive gene. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: CO2 (relating to SMD7561, SMD 7562, SMD 7261, SMD 7263, SMD 3710, SMD 4649, SMD 4650)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

CO<sub>2</sub> genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### CO2 Genes Identified By Cluster Analyses Of Differential Expression



CO2 Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of CO2 genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID CO2 (relating to SMD7561, SMD 7562, SMD 7261, SMD 7263, SMD 3710, SMD 4649, SMD 4650) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

CO2 Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of CO2 genes. A group in the MA\_clust is considered a CO2 pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

CO2 Genes Identified By Amino Acid Sequence Similarity

CO2 genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis CO2 genes. Groups of CO2 genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a CO2 pathway or network is a group of proteins that also exhibits CO2 functions/utilities.

III.D.1.USE OF CO2 RESPONSIVE GENES TO MODULATE PHENOTYPES

CO2 responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Catabolism

- Energy Generation, ATP, etc.
- Metabolism
- Carbohydrate Synthesis
- Growth Rate
- 5 - Whole Plant, including Height, Flowering Time, etc.
- Organs
- Flowers
- Fruits
- Stems
- 10 - Leaves
- Roots
- Lateral Roots
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation;
- 15 - Number, Size, and Weight of
- Flowers;
- Seeds;
- Branches;
- Leaves;
- 20 - Total Plant Nitrogen Content
- Amino Acid/Protein Content of Whole Plant or Parts
- Seed Yield
  - Number, Size, Weight, Harvest Index
  - Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, and
  - 25 Carbohydrate
- Fruit Yield
- Number, Size, Weight, Harvest Index
- Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, Carbohydrate, Water
- 30 - Photosynthesis
  - Carbon Dioxide Fixation

To improve any of the phenotype(s) above, activities of one or more of the CO<sub>2</sub> responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a  
5 plant can be transformed according to Bechtold and Pelletier (1998, *Methods. Mol. Biol.* 82:259-266) and/or screened for variants as in Winkler et al. (1998) *Plant Physiol* 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Saito et al. (1994, *Plant Physiol.* 106: 887-95), Takahashi et al (1997, *Proc. Natl. Acad. Sci. USA* 94: 11102-07) and Koprivova et al. (2000, *Plant Physiol.* 122: 737-46).

### III.D.2.USE OF CO<sub>2</sub> RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

5      The activities of one or more of the CO<sub>2</sub> responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Division	Cell Cycle Control Genes	Masle (2000) Plant Physiol. 122: 1399-1415
Starch Biosynthesis	Starch Biosynthesis Enzymes And Pathways	Ludewig et al., (1998) FEBS Lett. 429: 147-151
Photosynthesis	Photosynthetic Enzymes, e.g., Rubisco	Cheng et al., (1998) Plant Physiol 166: 715-723
Respiration	Energy Metabolism Pathways	Musgrave et al., (1986) Proc. Natl. Acad. Sci. USA 83: 8157-8161
CO <sub>2</sub> Uptake	Guard Cell Stomata Control Systems	Allen et al., Plant Cell (1999) 11(9): 1785-1798 Ichida et al., Plant Cell (1997) 9(10): 1843-1857 Hedrich et al., EMBO J (1993) 12(3): 897-901
Coordination Of Carbon And Nitrogen Metabolism	Light-Regulation Of Major Central Carbon And Nitrogen Metabolic Pathways To Coordinate Growth	Lam et al. (1998) Plant J. 16(3): 345-353 Lejay et al. (1999) Plant J. 18(5): 509-519; and Oliveira et al. (1999) Plant.

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		Phys. 121: 301-309
	Carbohydrate And Nitrogen Control Of Carbohydrate And Organic Nitrogen Accumulation Pathways	Lam et al. (1998) supra; Lejay et al. (1999) supra; and Oliveira et al. (1999) supra

Other biological activities that can be modulated by the CO<sub>2</sub> responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

CO<sub>2</sub> responsive genes are characteristically differentially transcribed in response to fluctuating CO<sub>2</sub> levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables report the changes in transcript levels of various CO<sub>2</sub> responsive genes that are differentially expressed in response to high CO<sub>2</sub> levels.

Profiles of these different CO<sub>2</sub> responsive genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts	Responders To Higher Levels Of CO <sub>2</sub>	<ul style="list-style-type: none"> <li>Changes In Generation Of ATP</li> <li>Changes In Catabolism And Anabolism Enzymes and Pathways</li> </ul>	<ul style="list-style-type: none"> <li>Transporters</li> <li>Catabolic And Anabolic Enzymes</li> <li>Change In Cell Membrane Structure</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	Genes Induced By CO <sub>2</sub>	<ul style="list-style-type: none"> <li>• Activation Of Krebs Cycle</li> <li>• Specific Gene Transcription Initiation</li> <li>• Changes In Carbohydrate Synthesis</li> <li>• Changes In Chloroplast Structure</li> <li>• Changes In Photosynthesis</li> <li>• Changes In Respiration</li> </ul>	<p>And Potential</p> <ul style="list-style-type: none"> <li>• Kinases And Phosphatases</li> <li>• Transcription Activators And Repressors</li> <li>• Change In Chromatin Structure And/Or Localized DNA Topology</li> <li>• Redox Control</li> </ul>
Down-Regulated Transcripts	<p>Responders To Higher Levels Of CO<sub>2</sub></p> <p>Genes Repressed By CO<sub>2</sub></p>	<ul style="list-style-type: none"> <li>• Changes In Pathways And Processes Operating In Cells</li> <li>• Changes In Catabolism and Anabolism</li> <li>• Changes in Chloroplast Structure</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Change In Protein Structure By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>• Change In Chromatin Structure And/Or DNA Topology</li> <li>• Stability Of Factors For Protein Synthesis And Degradation</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
			<ul style="list-style-type: none"> <li>Metabolic Enzymes</li> </ul>

#### Use of Promoters of CO<sub>2</sub> responsive Genes

Promoters of CO<sub>2</sub> responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the CO<sub>2</sub> responsive genes where the desired sequence is operably linked to a promoter of a CO<sub>2</sub> responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

III.E. MITOCHONDRIA ELECTRON TRANSPORT (RESPIRATION) GENES,  
GENE COMPONENTS AND PRODUCTS

One means to alter flux through metabolic pathways is to alter the levels of proteins in  
the pathways. Plant mitochondria contain many proteins involved in various metabolic  
processes, including the TCA cycle, respiration, and photorespiration and particularly the  
electron transport chain (mtETC). Most mtETC complexes consist of nuclearly-encoded  
mitochondrial proteins (NEMPs) and mitochondrially-encoded mitochondrial proteins (MEMPs).  
NEMPs are produced in coordination with MEMPs of the same complex and pathway and with  
other proteins in multi- organelle pathways. Enzymes involved in photorespiration, for example,  
are located in chloroplasts, mitochondria, and peroxisomes and many of the proteins are  
nuclearly-encoded. Manipulation of the coordination of protein levels within and between  
organelles can have critical and global consequences to the growth and yield of a plant. Genes  
which are manipulated by interfering with the mtETC have been characterized using microarray  
technology.

Microarray technology allows monitoring of gene expression levels for thousands of  
genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to  
glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US  
Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from  
such microarray experiments conducted with AFGC chips containing about 10,000 non-  
redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA  
of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in the  
presence of the ETC inhibitor, 10 mM antimycin A compared with the control lacking antimycin  
A, were identified, compared to the Ceres full length cDNA and genomic sequence databanks,  
and equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or  
AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres  
clones which are up or down regulated over controls, thereby indicating the Ceres clones that  
represent respiration responsive genes.

Examples of genes and gene products that are responsive to antimycin A block of  
respiration are shown in the Reference, Sequence, Protein Group, Protein Group Matrix,



MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. While respiration responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different respiration responsive polynucleotides and/or gene products that have similar  
5 transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways. Whole  
10 pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. Manipulation of one or more respiration responsive gene activities are useful to modulate the biological processes and/or phenotypes listed below.

Such respiration responsive genes and gene products can function to either increase or dampen the phenotypes or activities below. Further, promoters of respiration responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by respiration or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the respiration responsive genes when the desired sequence is operably linked to a promoter of a respiration responsive gene. The MA\_diff and/or AFLP\_diff MA\_diff and/or  
20 AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Mitochondria-Electron Transport (relating to SMD 8061, SMD 8063)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental  
25 detail see the Example section below.

Mitochondria-Electron Transport genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

30 Mitochondria-Electron Transport Genes Identified By Cluster Analyses Of Differential Expression

Mitochondria-Electron Transport Genes Identified By Correlation To Genes That  
Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Mitochondria-Electron Transport genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Mitochondria-Electron Transport (relating to SMD 8061, SMD 8063) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Mitochondria-Electron Transport Genes Identified By Correlation To Genes That  
Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Mitochondria-Electron Transport genes. A group in the MA\_clust is considered a Mitochondria-Electron Transport pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Mitochondria-Electron Transport Genes Identified By Amino Acid Sequence Similarity

Mitochondria-Electron Transport genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Mitochondria-Electron Transport genes. Groups of Mitochondria-Electron Transport genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Mitochondria-Electron Transport pathway or network is a group of proteins that also exhibits Mitochondria-Electron Transport functions/utilities.

III.E.1. USE OF RESPIRATION RESPONSIVE GENES TO MODULATE  
PHENOTYPES

Respiration responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Catabolism
- Energy Generation, ATP, etc.
- Growth Rate
- Whole Plant, including Height, Flowering Time, etc.
- Organs
- Flowers
- Fruits
- Stems
- Leaves
- Roots
- Lateral Roots
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation;
- Number, Size, and Weight of
- Flowers;
- Seeds;
- Branches;
- Leaves;
- Total Plant Nitrogen Content
- Amino Acid/Protein Content of Whole Plant or Parts
- Seed Yield
  - Number, Size, Weight, Harvest Index
  - Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, and Carbohydrate
- Fruit Yield
- Number, Size, Weight, Harvest Index

- Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, Carbohydrate, Water
- Photosynthesis
  - Carbon dioxide fixation

To improve any of the phenotype(s) above, activities of one or more of the respiration responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Saito et al. (1994, Plant Physiol. 106: 887-95), Takahashi et al (1997, Proc. Natl. Acad. Sci. USA 94: 11102-07) and Koprivova et al. (2000, Plant Physiol. 122: 737-46).

#### III.E.2. USE OF RESPIRATION-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the respiration responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Respiration and energy-related processes	Mitochondrial Electron Transport Chain	Passam et al. (1973) Biochem Biophys. Acta 325: 54-61
	Alternative oxidase pathway	Saisho et al. (1997) Plant Mol. Biol. 35: 585-600 Vanlerberghe and McIntosh (1994) Plant Physiol. 105: 867-874

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	ATP generation pathways ATP utilization pathways	Mahler and Cordes (1966) In Biological Chemistry, Harper and Row
	Chloroplast energy related pathways	Foyer et al. (1989) Arch. Biochem. Biophys. 268: 687-697 Mills et al. (1978) Biochem. Biophys. Acta 504: 298-309
	Peroxisome energy related pathways	Olsen (1998) Plant mol. Biol. 38: 163-89
	Cytoplasmic energy related pathways	Roberts et al. (1995) Febs Letters 373: 307-309
	Catabolism and Anabolism	Mahler and Cordes (1966) In Biological Chemistry, Harper and Row
	Aerobic versus anaerobic pathways	Mahler and Cordes (1966) In Biological Chemistry, Harper and Row
Coordination of Carbon and Nitrogen Metabolism	Light-regulation of major central carbon and nitrogen Metabolic pathways to coordinate growth	Lam et al. (1998) Plant J. 16(3): 345-353 Lejay et al. (1999) Plant J. 18(5): 509-519; and Oliveira et al. (1999) Plant. Phys. 121: 301-309
	Carbohydrate and nitrogen control of carbohydrate and organic nitrogen accumulation	Lam et al. (1998) Plant J. 16(3): 345-353 Lejay et al. (1999) Plant J. 18(5):

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	pathways	509-519; and Oliveira et al. (1999) Plant. Phys. 121: 301-309

Other biological activities that can be modulated by the respiration genes and gene products are listed in the REF Tables. Assays for detecting such biological activities are described in the Protein Domain table.

Respiration responsive genes are differentially expressed in response to inhibition of mitochondrial electron transport by antimycin A. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various respiration responsive genes that are differentially expressed in response to this treatment.

Profiles of these different respiration genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	<p>Responders to inhibition of mitochondrial electron transport respiration</p> <p>Genes induced by inhibition of mitochondrial electron transport</p>	<ul style="list-style-type: none"> <li>• Changes in generation of ATP</li> <li>• Alternate oxidase induction</li> <li>• Changes in catabolic and anabolic enzymes and pathways</li> <li>• Specific gene transcription initiation</li> <li>• Changes in electron transport proteins</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Catabolic and anabolic enzymes</li> <li>• Changes in cell and organelle membrane structures and potentials</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> <li>• Redox control</li> </ul>
Down-regulated transcripts	<p>Responders to inhibition of mitochondrial electron transport</p> <p>Genes repressed by inhibition of mitochondrial</p>	<ul style="list-style-type: none"> <li>• Changes in ATP generating pathways</li> <li>• Changes in pathways and processes operating in cells</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	electron transport	<ul style="list-style-type: none"> <li>• Induction of aerobic pathways</li> <li>• Changes in catabolism and anabolism</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Catabolic and anabolic enzymes</li> <li>• Changes in cell and organelle membrane structures and potentials</li> <li>• Change in chromatin structure and/or localized DNA topology</li> <li>• changes</li> <li>• Stability factors for protein synthesis and degradation</li> <li>• Metabolic enzymes</li> </ul>
		<ul style="list-style-type: none"> <li>• Changes in redox activities</li> </ul>	<ul style="list-style-type: none"> <li>• Changes in redox enzymes</li> </ul>

#### Use of Promoters of Respiration Genes

Promoters of Respiration genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Respiration genes where the desired



sequence is operably linked to a promoter of a Respiration gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

III.F. PROTEIN DEGRADATION GENES, GENE COMPONENTS AND  
PRODUCTS

One of the components of molecular mechanisms that operate to support plant development is the "removal" of a gene product from a particular developmental circuit once the substrate protein is not functionally relevant anymore in temporal and/or spatial contexts. The "removal" mechanisms can be accomplished either by protein inactivation (e.g., phosphorylation or protein-protein interaction) or protein degradation most notably via ubiquitination-proteasome pathway. The ubiquitination-proteasome pathway is responsible for the degradation of a plethora of proteins involved in cell cycle, cell division, transcription, and signal transduction, all of which are required for normal cellular functions. Ubiquitination occurs through the activity of ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin-protein ligases (E3), which act sequentially to catalyze the attachment of ubiquitin (or other modifying molecules that are related to ubiquitin) to substrate proteins (Hochstrasser 2000, Science 289: 563). Ubiquitinated proteins are then routed to proteasomes for degradation processing [2000, Biochemistry and Molecular Biology of Plants, Buchanan, Gruissem, and Russel (eds), Amer. Soc. of Plant Physiologists, Rockville, MD]. The degradation mechanism can be selective and specific to the concerned target protein (Joazeiro and Hunter 2001, Science 289: 2061; Sakamoto et al., 2001, PNAS Online 141230798). This selectivity and specificity may be one of the ways that the activity of gene products is modulated.

III.F.1. IDENTIFICATION OF PROTEIN DEGRADATION GENES, GENE  
COMPONENTS AND PRODUCTS

"Protein degradation" genes identified herein are defined as genes, gene components and products associated with or dependant on the ubiquitination – proteasome protein degradation process. . Examples of such "protein degradation" genes and gene products are shown in the Reference and Sequence Tables. The biochemical functions of the protein products of many of these genes are also given in the Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. Selected genes, gene components and gene products of the invention can be used to modulate many plant traits from architecture to yield to stress tolerance.

"Protein Degradation" Genes, Gene Components And Products Identified By Phenotypic Observations

5 "Protein degradation" genes herein were discovered and characterized from a much larger set of genes in experiments designed to find the genes associated with the increased number of lateral branches (and secondary inflorescences) that are formed per cauline node. In these experiments, "protein degradation" genes were identified using a mutant with these characteristics. The gene causing the changes was identified from the mutant gene carrying an inserted tag. The mutant plant was named 13B12-1 and the mutant was in the E2 conjugating enzyme gene of the ubiquitination process. Compared to "wild-type" parental plants, the mutant plants exhibited multiple lateral stems per node and multi-pistillated flowers. For more experimental detail, see Example section below.

10  
15 Protein Degradation Genes, Gene Components And Products Identified By Differential Expression

"Protein degradation" genes were also identified by measuring the relative levels of mRNA products in the mutant plant 13B12-1 lacking the E2 conjugating enzyme versus a "wild-type" parental plant. Specifically, mRNAs were isolated from 13B12-1 and compared with mRNAs isolated from wild-type plants utilizing microarray procedures. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108451). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

20  
25 Protein Degradation genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

Protein Degradation Genes Identified By Cluster Analyses Of Differential Expression

30 Protein Degradation Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Protein Degradation genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108451 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

#### Protein Degradation Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Protein Degradation genes. A group in the MA\_clust is considered a Protein Degradation pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Protein Degradation Genes Identified By Amino Acid Sequence Similarity

Protein Degradation genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Protein Degradation genes. Groups of Protein Degradation genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Protein Degradation pathway or network is a group of proteins that also exhibits Protein Degradation functions/utilities.

These differentially expressed genes include genes associated with the degradation process and the genes whose expression is disturbed by the aberrant ubiquitination.

Examples of phenotypes, biochemical activities, and transcription profiles that can be modulated using these genes, gene components and gene products are described above and below.

III.F.2. USE OF "PROTEIN DEGRADATION" GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE PHENOTYPES

The "protein degradation" genes, their components and products of the instant invention are useful for modulating one or more processes required for post-translational modification (e.g., ubiquitination) and degradation or inactivation of substrate proteins and also the pathways and processes that are associated with protein inactivation that are important for either or all of the following: (I) cell proliferation; (II) cell differentiation; and (III) cell death. The "protein degradation" genes, gene components and gene products are useful to alter or modulate one or more of the following phenotypes:

I. Cell Proliferation

A. Cell properties

Cell properties can be critically altered by the maintenance or not of regulatory proteins

1. Cell size
2. Cell division, rate and direction
3. Cell elongation

B. Plant size

The following parts of a plant can be modulated by "protein degradation" genes, gene components or gene products to affect plant size:

1. Roots
  - (a) Primary
  - (b) Lateral
  - (c) Root hairs
  - (d) Root cap
  - (e) Apical meristem
  - (f) Epidermis
  - (g) Cortex

- (h) Stele
- 2. Stem
  - (a) Phloem
  - (b) Xylem
  - (c) Nodes
  - (d) Internodes
  - (e) Leaves
  - (f) Shoot apical meristem
  - (g) Cauline
  - (h) Rosette
  - (i) Petioles
- 3. Flowers
  - (a) Receptacle
  - (b) Sepals, Petals, and Tepals
  - (c) Androecium
  - (d) Stamen
  - (e) Anther
  - (f) Pollen
  - (g) Filament
  - (h) Gynoecium
  - (i) Carpel
  - (j) Ovary
  - (k) Style
  - (l) Stigma
  - (m) Ovule
  - (n) Pedicel and Peduncle
- 4. Seeds
  - (a) Placenta
  - (b) Embryo
  - (c) Cotyledon
  - (d) Endosperm

- (e) Suspensor
- (f) Seed coat (testa)
- 5. Fruits
  - (a) Pericarp – thickness, texture
  - (b) Exocarp
  - (c) Mesocarp
  - (d) Encocarp

## II. CELL DIFFERENTIATION

The intracellular levels of many proteins are regulated by ubiquitin-proteasome proteolysis. Without proper regulation of protein levels, normal cell differentiation can be altered. Examples of cell differentiation and development that are useful to modulated by the genes and gene products of this invention are as follows:

### A. Roots

The polynucleotides and polypeptides of this invention can be used to control root structure and function. Examples are as follows:

- 1. Size
  - (a) Length of primary roots
  - (b) Length of lateral roots
- 2. Function (for more detail see Root section)
- B. Branching and stem formation
  - 1. Multiple pistils
  - 2. Multiple lateral stems or secondary inflorescence per cauline node
  - 3. Internode length
- C. Cell differentiation and/or development in response to hormones
  - 1. Auxin

## III. CELL DEATH

Programmed cell death can result from specific and targeted degradation of critical substrate proteins (e.g., transcription factors, enzymes, and proteins involved in signal transduction). Thus, alteration of "protein degradation" genes, their gene products, and the

corresponding substrate proteins that they are acting upon are useful to modulate the vigor and yield of the plant overall as well as distinct cells, organs, or tissues. Examples of traits that can be modulated by these genes and gene products include:

- A. Sterility or reproduction
- B. Seedling lethality

Uses Of Plants That Are Modified As Described Above

Genes that control fundamental steps in regulatory pathways, such as protein inactivation, that in turn influence cascades and networks of other genes and processes are extremely useful. They and their component parts can be used selectively to manipulate development in specific cells, tissues and organs, including meristems when genes are designed to inactivate the normal genes only in specific cells, tissues and organs or to promote protein production where it is not normally produced. They can also be used to promote/control cell death.

Other "protein degradation" genes described here are components of the pathways determining organ identity and phenotypes. These and their component parts are also useful for modifying the characteristics of specific cells, tissues and organs when regulated appropriately. Thus "protein degradation" genes have wide utility for achieving the following:

- A. Better plant survival by decreased lodging
- B. Better responses to high plant density
- C. Better stress tolerance
- D. Better animal (including human) nutrition values
- E. Improved dietary mineral nutrition
- F. More vigor, growth rate and yield
  - 1. Growth rate
    - (a) Whole plant, including height, flowering time, branching, etc.
    - (b) Seedling
    - (c) Coleoptile elongation
    - (d) Young leaves
    - (e) Flowers
    - (f) Seeds
    - (g) Fruit



2. Yield

(a) Biomass

- Fresh and dry weight during any time in plant life, including maturation and senescence

(b) Root/tuber yield

- Number, size, weight, harvest index

- Content and composition, e.g. amino acid, jasmonate, oil, protein and starch

(c) Number of flowers

(d) Seed yield

- Number, size, weight, harvest index
- Content and composition, e.g. amino acid, jasmonate, oil, protein and starch

(e) Fruit yield

- Number, size, weight, harvest index, post harvest quality
- Content and composition, e.g. amino acid, jasmonate, oil, protein and starch

To regulate any of the phenotype(s) above, activities of one or more of the "protein degradation" genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. In addition, a synthetic molecule containing specific domains from "protein degradation" genes or gene product and/or in combination with other domains from gene products that are not necessarily related to protein degradation pathway can be constructed to target the degradation or inactivation of specific substrate proteins. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Dolan et al. (1993, Development 119: 71-84), Dolan et al. (1997, Development 124: 1789-98), Crawford and Glass (1998, Trends Plant Science 3: 389-95), Wang et al. (1998, PNAS USA 95: 15134-39), Gaxiola et al. (1998, PNAS USA 95: 4046-50), Apse et al. (1999, Science 285: 1256-58), Fisher and Long (1992, Nature

357: 655-60), Schneider et al. (1998, Genes Devel 12: 2013-21) and Hirsch (1999, Curr Opin Plant Biol. 2: 320-326).

III.F.3. USE OF PROTEIN DEGRADATION GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

One or more of the "protein degradation" genes and their components can be used to  
modulate biochemical or metabolic activities, processes and/or pathways such as those noted  
below. Such biological activities can be measured according to the citations included in the  
Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth, Differentiation and Development	Auxin response	Schwechheimer et al, Science 292: 1379 (2001); Leyser et al, Nature 8: 161 (1993); Lasswell et al, Plant Cell 12: 2395 (2000)
	Photomorphogenesis via leaf cells and meristems	Schwechheimer et al, Science 292: 1379 (2001)
	Apical dominance via shoot meristems	Schwechheimer et al, Science 292: 1379 (2001)
	Lateral root development via root meristem	Xie et al, Genes Dev 14: 3024 (2000)
	Hypocotyl, shoot elongation by hormone controlled process	Nagpal et al, Plant Physiol 123: 563 (2000)
Gene Expression and related cellular processes	mRNA stability	Johnson et al, PNAS 97: 13991 (2000);
	Gene activation	Pham and Sauer, 289: 2357 (2000)
	Cell division and cell cycle control in meristems	King et al, Cell 81: 279 (1995); Ciechanover et al, Cell 37: 57 (1984); Finley et al, Cell 37: 43 (1984);

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		Robzyk et al, Science 287: 501 (2000)
	Chromatin remodeling	Roest et al, Cell 86: 799 (1996)
	Post-translational modification and organelle targeting of proteins	Biederer et al, Science 278: 1806 (1997)

Other biological activities that can be modulated by the "protein degradation" gene, gene components and products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

#### III.F.4. USE OF PROTEIN DEGRADATION GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE TRANSCRIPTION LEVELS OF OTHER GENES

The expression of many genes is "up regulated" or "down regulated" in the 13B12-1 mutant because some protein degradation genes and their products are integrated into complex networks that regulate transcription of many other genes. Some protein degradation genes are therefore useful for modifying the transcription of other genes and hence complex phenotypes, as described above. Profiles of "protein degradation" genes are described in the Table below with associated biological activities. "Up-regulated" profiles are those where the gene produces mRNA levels that are higher in the 13B12-1 as compared to wild-type plant; and vice-versa for "down-regulated" profiles.

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
Up Regulated Transcripts	<ul style="list-style-type: none"> <li>Genes induced as a consequence of mutant ubiquitination degradation system</li> <li>Genes repressed by "protein degradation" system directly or indirectly</li> <li>Genes repressed or mRNAs degraded as a consequence of mutant ubiquitination degradation process</li> </ul>	<ul style="list-style-type: none"> <li>Shoot formation</li> <li>Lateral stem, lateral and main inflorescence development</li> <li>Internode elongation</li> <li>Node determination and development</li> <li>Root formation</li> <li>Lateral root development</li> <li>Proper response to Auxin and other growth regulators</li> <li>Seed dormancy and seed development</li> <li>Resistance to drought and other forms of stress</li> <li>Secondary metabolite biosynthesis</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Activators and Repressors</li> <li>Chromatin Structure and/or Localized DNA Topology determining proteins</li> <li>Methylated DNA binding proteins</li> <li>Kinases, Phosphatases</li> <li>Signal transduction pathway proteins</li> <li>Transporters</li> <li>Metabolic Enzymes</li> <li>Cell cycle checkpoint proteins</li> <li>Cell Membrane Structure And Proteins</li> <li>Cell Wall Proteins</li> <li>Proteins involved in secondary metabolism</li> <li>Seed storage</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
			metabolism
Down Regulated Transcripts	<ul style="list-style-type: none"> <li>Genes activated by "protein degradation" systems directly or indirectly</li> </ul>		

"Protein degradation" genes and gene products can be modulated alone or in combination as described in the introduction. Of particular interest are combination of these genes and gene products with those that modulate hormone responses and/or metabolism. Hormone responsive and metabolism genes and gene products are described in more detail in the sections above. Such modification can lead to major changes in plant architecture and yield.

#### Use Of Promoters And "Protein Degradation Genes, Gene Components And Products"

Promoters of "protein degradation" genes, as described in the Reference tables, for example, can be used to modulate transcription of any polynucleotide, plant or non plant to achieve synthesis of a protein in association with production of the ubiquitination -proteasome pathway or the various cellular systems associated with it. Additionally such promoters can be used to synthesize antisense RNA copies of any gene to reduce the amount of protein product produced, or to synthesize RNA copies that reduce protein formation by RNA interference. Such modifications can make phenotypic changes and produce altered plants as described above.

III.G. CAROTENOGENESIS RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Carotenoids serve important biochemical functions in both plants and animals. In plants, carotenoids function as accessory light harvesting pigments for photosynthesis and to protect chloroplasts and photosystem II from heat and oxidative damage by dissipating energy and scavenging oxygen radicals produced by high light intensities and other oxidative stresses. Decreases in yield frequently occur as a result of light stress and oxidative stress in the normal growth ranges of crop species. In addition light stress limits the geographic range of many crop species. Modest increases in oxidative stress tolerance would greatly improve the performance and growth range of many crop species. The development of genotypes with increased tolerance to light and oxidative stress would provide a more reliable means to minimize crop losses and diminish the use of energy-costly practices to modify the soil environment.

In animals carotenoids such as beta-carotene are essential provitamins required for proper visual development and function. In addition, their antioxidative properties are also thought to provide valuable protection from diseases such as cancer. Modest increases in carotenoid levels in crop species could produce a dramatic effect on plant nutritional quality. The development of genotypes with increased carotenoid content would provide a more reliable and effective nutritional source of Vitamin A and other carotenoid derived antioxidants than through the use of costly nutritional supplements.

Genetic changes produced through DNA mutation in a plant can result in the modulation of many genes and gene products. Examples of such mutation altered genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor, nutritional content and seed yield.

While carotenoid synthesis and/or oxidative stress responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different carotenoid biosynthetic polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of an carotenoid synthesis or oxidative stress protective polynucleotide and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress pathways, nutritional pathways and

development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in a common pathway.

Such carotenoid synthesis/oxidative stress tolerance genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in light intensity or in the absence of osmotic fluctuations. They were discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA products participate in carotenogenesis. These experiments made use of an Arabidopsis mutant (Or) having an accumulation of up to 500 times more beta-carotene than wild-type in non-photosynthetic tissues.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The USArabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing some 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in the mutant plant compared with wild type seedlings were identified, compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which represent Carotenoid synthesis/oxidative stress tolerance responsive genes. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Cauliflower (relating to SMD 5329, SMD 5330)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Carotenogenesis genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication



5

Carotenogenesis Genes Identified By Cluster Analyses Of Differential Expression

Carotenogenesis Genes Identified By Correlation To Genes That Are  
Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Carotenogenesis genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Cauliflower (relating to SMD 5329, SMD 5330) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Carotenogenesis Genes Identified By Correlation To Genes That Cause  
Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Carotenogenesis genes. A group in the MA\_clust is considered a Carotenogenesis pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Carotenogenesis Genes Identified By Amino Acid Sequence Similarity

Carotenogenesis genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Carotenogenesis genes. Groups of Carotenogenesis genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a

Carotenogenesis pathway or network is a group of proteins that also exhibits Carotenogenesis functions/utilities.

5

10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995

III.G.1.USE OF CAROTENOID SYNTHESIS,/OXIDATIVE STRESS  
TOLERANCE RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE PHENOTYPES

Carotenoid synthesis/oxidative stress tolerance genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth rate
- Whole Plant, including Height, Flowering Time, etc.
- Seedling
- Organ
  - Stem
  - Leaves
  - Roots
  - Flowers
  - Fruits
  - Seeds
  - Yield
  - Size, Weight
- Seed Development
- Embryo
- Germination
- Cell Differentiation
- Chloroplasts

- Plant nutrition
- Uptake and assimilation of organic compounds
- Uptake and assimilation of inorganic compounds
- Animal (including human) nutrition
- Improved dietary mineral nutrition
- Stress Responses
- Drought
- Cold
- Osmotic

To improve any of the phenotype(s) above, activities of one or more of the Carotenoid synthesis/oxidative stress tolerance genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Friedrich, (1999, JAMA 282: 1508), Kumar et al. (1999, Phytochemistry 51: 847-51), La Rocca et al. (2000, Physiologia Plantarum 109: 51-7) and Bartley (1994, In: Ann Rev Plant Physiol Plant Molec Biol, Jones and Somerville, eds, Annual Reviews Inc, Palo Alto, CA).

### III.G.2. USE OF CAROTENOID SYNTHESIS/OXIDATIVE STRESS TOLERANCE RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the carotenoid synthesis/oxidative stress tolerance genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation and Development	Chloroplast biosynthesis	Kumar et al. (1999) Phytochemistry 51: 847-51 Fraser et al. (1994) Plant Physiol 105: 405-13
Metabolism	Carotenoid biosynthesis	Kumar et al. (1999) Phytochemistry 51: 847-51
	Herbicide resistance	La Rocca et al. (2000) Physiologia Plantarum 109: 51-57
	Regulate abscisic acid levels	Tan et al. (1997) PNAS USA 94: 12235-40
	Drought, cold and osmotic tolerance	Tan et al. (1997) PNAS USA 94: 12235-40

Other biological activities that can be modulated by the Carotenoid synthesis, oxidative stress tolerance genes and gene products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Protein Domain table.

5 Profiles of these different carotenoid synthesis/oxidative stress tolerance responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	Genes induced during carotenoid synthesis/oxidative stress tolerance activity	<ul style="list-style-type: none"> <li>• Gene Repression/Induction activity</li> <li>• Cell cycle progression</li> <li>• Chromatin condensation</li> <li>• Synthesis of metabolites and/or proteins</li> <li>• Modulation of transduction pathways</li> <li>• Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic enzymes</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> </ul>
Down-regulated transcripts	<p>Genes repressed during carotenoid synthesis/oxidative stress tolerance activity</p> <p>Genes with discontinued expression or unstable mRNA in conditions of reduced carotenoid synthesis/oxidative stress tolerance</p>	<ul style="list-style-type: none"> <li>• Gene repression/induction activity</li> <li>• Changes in pathways and processes operating in cells</li> <li>• Changes in metabolism other than carotenoid synthesis/oxidative stress tolerance</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> <li>• Change in chromatin structure and/or DNA topology</li> <li>• Stability of</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
			<p>factors for protein synthesis and degradation</p> <ul style="list-style-type: none"> <li>• Metabolic enzymes</li> </ul>

#### Use of Promoters of Carotenogenesis Responsive Genes

Promoters of Carotenogenesis responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Carotenogenesis responsive genes where the desired sequence is operably linked to a promoter of a Carotenogenesis responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

#### IV. VIABILITY GENES, GENE COMPONENTS AND PRODUCTS

##### IV.A. VIABILITY GENES, GENE COMPONENTS AND PRODUCTS

Plants contain many proteins and pathways that when blocked or induced lead to cell, organ or whole plant death. Gene variants that influence these pathways can have profound effects on plant survival, vigor and performance. The critical pathways include those concerned with metabolism and development or protection against stresses, diseases and pests. They also include those involved in apoptosis and necrosis. The applicants have elucidated many such genes and pathways by discovering genes that when inactivated lead to cell or plant death.

Herbicides are, by definition, chemicals that cause death of tissues, organs and whole plants. The genes and pathways that are activated or inactivated by herbicides include those that cause cell death as well as those that function to provide protection. The applicants have elucidated these genes.

The genes defined in this section have many uses including manipulating which cells, tissues and organs are selectively killed, which are protected, making plants resistant to herbicides, discovering new herbicides and making plants resistant to various stresses.

##### IV.A.1. IDENTIFICATION OF VIABILITY GENES, GENE COMPONENTS AND PRODUCTS

Viability genes identified here are defined as genes, gene components and products capable of inhibiting cell, tissue, organ or whole plant death or protecting cells, organs and plants against death and toxic chemicals or stresses. Examples of such viability genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff, MA\_clust, Knock-in and Knock-out tables. The biochemical functions of the protein products of many of these genes determined from comparisons with known proteins are also given in the Reference tables.

##### Viability Genes, Gene Components And Products Identified By Phenotypic Observations

These genes were discovered and characterized from a much larger set of genes by experiments designed to find genes that cause serious disturbances in progeny survival, seed germination, development, embryo and/or seedling growth. In these experiments, viability genes

were identified by either (1) ectopic expression of a cDNA in a plant or (2) mutagenesis of a plant genome. The plants were then cultivated and one or more of the following phenotypes, which varied from the parental wild-type was observed:

- A. Gametophytic loss of progeny seedlings (detected from a parent on the basis of a linked herbicide resistance gene showing abnormal segregation ratios, as revealed by treating with herbicide)
- B. Embryo death, resulting in some cases to loss of seed
- C. Pigment variation in cotyledons and leaves, including absence of chlorophyll, which leads to seedling death.
  - 1. Abinos
  - 2. Yellow/greens
- D. Cotyledons produced but no or few leaves and followed by seedling death.
- E. Very small plantlets

The genes identified in these experiments are shown in Tables X.

#### Viability Genes, Gene Components And Products Identified By Differential Expression

Viability genes were also identified from a much larger set of genes by experiments designed to find genes whose mRNA products changed in concentration in response to applications of different herbicides to plants. Viability genes are characteristically differentially transcribed in response to fluctuating herbicide levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table reports the changes in transcript levels of various viability genes in entire seedlings at 0, 4, 8, 12, 24, and 48 hours after a plant was sprayed with a Hoagland's nutrient solution enriched with either 2,4 D (Trimec), Glean, Grassgetter, Roundup, or Finale herbicides as compared to seedlings sprayed with Hoagland's solution only.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108467, 107871, 107876, 108468, 107881, 108465, 107896, 108466, 107886, 107891, 108501). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example



section below.

Viability genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a “+” or “-” indication

#### Viability Genes Identified By Cluster Analyses Of Differential Expression

##### Viability Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Viability genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108467, 107871, 107876, 108468, 107881, 108465, 107896, 108466, 107886, 107891, 108501 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Viability Genes Identified By Correlation To Genes That Cause

##### Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Viability genes. A group in the MA\_clust is considered a Viability pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Viability Genes Identified By Amino Acid Sequence Similarity

Viability genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Viability genes. Groups of Viability genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Viability pathway or network is a group of proteins that also exhibits Viability functions/utilities.

It is assumed that those gene activity changes in response to the toxic herbicides are either responsible, directly or indirectly, for cell death or reflect activation of defense pathways. These genes are therefore useful for controlling plant viability.

Examples of phenotypes, biochemical activities, or transcript profiles that can be modulated using selected viability gene components are described above and below.

#### IV.A.2.USE OF VIABILITY GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE PHENOTYPES

Deficiencies in viability genes can cause cell death at various rates and under various conditions. Viability genes can be divided into two classes; (1) those that lead to cell death under permissive growth conditions and (2) those that cause cell demise under restrictive conditions. Examples of the first class are viability genes which encode toxins or which participate in the programmed cell death pathway(s). Disruption of metabolic pathways, such as amino acid synthesis, may not cause death when the cell is supplemented with appropriate amino acids, but can cause death under more restrictive conditions.

Some deficiencies in viability genes identified cause the organism as a whole to die, while other genes cause death only of a specific subset of cells or organs. For example, genes identified from embryo viability phenotypes can cause an entire organism to die. In contrast, genes characterized from gametophytic lethals may inhibit cell growth only in a select set of cells. In addition, some viability genes may not cause an immediate demise. A seedling lethal phenotype is one such example, where a seed germinates and produces cotyledons but the plant dies before producing any true leaves. Yellow-green pigment mutants provide yet another set of examples. In some cases, the plant produces a number of yellow-green leaves but dies before producing any seed, due in part, to the necessity to produce chlorophyll in functioning chloroplasts to fix CO<sub>2</sub>.

Viability genes, in which mutational deficiencies lead to death, carry no duplicates in the haploid plant genome. They thus may be especially likely to promote viability and vigor when expressed more optimally in a plant, in specific tissues or throughout the plant.

Proteins which lead to death when inactivated, and other proteins in the pathways in which they act, are potential targets for herbicides. In this kind of application, chemicals specifically capable of interacting with such proteins are discovered. Typically, this could be done by designing a gene involving the relevant viability gene, that also facilitates a rapid easily measured assay for the functioning of the protein product, and treating plants containing the new genes with the potential herbicides. Those chemicals specifically interfering with the protein activity can then easily be selected for further development.

Genes whose products interact directly with a herbicide can also be modified such that the herbicide no longer inactivates the protein. Such genes are useful for making herbicide resistant plants, valuable in agriculture.

Many of the genes activated or inactivated by the herbicides define genes involved in the pathways that protect the plant against damage and stresses. These genes and gene components, especially those regulating such pathways, are especially useful for enhancing the ability of plants to withstand specific stresses, including herbicides. [See the sections on Stress responsive genes, gene components and products.]

Genes that cause cellular death can be used to design new genes that cause death of specific cells and tissues and hence new valuable products. For example, activation of genes causing death in cells specifying seeds can be used to produce fruits lacking seeds. They can also be used to prevent cell death by pathogens and pests.

The genes and gene components of the instant invention are useful to modulate one or more processes that affect viability and vigor at the (I) cellular level; (II) organelle level; (III) organ level; or (IV) overall organism level.

Examples of phenotypes that are modulated are described above and below:

I. CELLULAR LEVEL:

Viability genes and gene products are useful in modulating cellular changes in:

- A. Cell size
- B. Cell differentiation
- C. Cell division
- D. Cell longevity
- E. Cell position

F. Cytotoxins

II. ORGANELLE LEVEL

The development, growth, and viability of chloroplasts and/or mitochondria can be  
5 modulated by the genes and gene products of the instant invention:

III. ORGAN LEVEL:

The invention is also useful to regulate the development, growth, and viability of the  
following organs:

- 10 A. Flower 1. number
- 2. size
- B. Seed
- 1. Size
- Number
- 5 Composition
- Amino Acid, carbohydrates, lipid, and secondary metabolites
- C. Fruit
- 1. Size
- 2. Number
- 20 3. Composition
- Amino Acid, carbohydrates, lipid, and secondary metabolites
- 1. Fruit Drop
- 2. Fruit Ripening
- D. Leaf
- 25 1. Size
- 2. Composition
- 3. Amino acid, carbohydrates, lipid, and secondary metabolites
- 4. Photoefficiency
- 5. Abscission
- 30 6. Senescence
- E. Stem

F. Root

#### IV. OVERALL ORGANISM LEVEL

The following traits can be modulated with the genes, gene components and gene products of this invention to affect the strength of a plant as a whole:

- A. Vigor
  - Increased biomass
- B. Stress Tolerance
  - 1. Cold
  - 2. Drought
  - 3. Heat
  - 4. Herbicide
  - 5. Oxidative
  - 6. Salt
- C. Pathogen Resistance

To regulate any of the phenotype(s) above, activities of one or more of the viability genes or gene products can be modulated in an organism and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (Methods, Mol. Biol. 82:259-266 (1998)) and/or screened for variants as in Winkler et al., Plant Physiol. 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed.

#### IV.A.3. USE OF VIABILITY GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The viability genes, their components and/or products can be used to modulate processes, biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Amino Acid Synthesis	Aceto -lactate synthase	Hershey et al. (1999) Plant Mol. Biol. 40, 795-806
Cell Wall Synthesis	Cellulose synthase	Peng et al. (2001) Plant Physiol. 126, 981-982 Kawagoe and Delmer (1997) Genet Eng. 19, 63-87
Nucleotide Synthesis	Coenzyme A biosynthesis	Kupke et al. (2001) J. Biol. Chem. 276, 19190-19196
Lipid Synthesis	Oleosin biosynthesis	Singh et al. (2000) Biochem. Soc. Trans. 28, 925-927 Zou et al. (1996). Plant Mol. Biol. 31, 429-433
Hormone Signaling Pathways	Brassinolide and light signal transduction	Kang et al. (2001) Cell 105, 625-636
Hormone Biosynthesis	Cytokinin biosynthesis	Takei et al, (2001) J. Biol. Chem. 276, 26405-26410
Secondary Metabolites	• Carotenoid biosynthesis	Estevez et al. (2001) J. Biol. Chem. 276, 22901-22909 Carol and Kuntz (2001) Trendy Plant Sci. 6, 31-36 Pogson and Rissler (2001) Phil. Trans. Roy. Soc. Lord. B 355, 1395-1400
Clearing of Toxic Substances	Ubiquitination	
Growth, Differentiation	• Farnesylation	Pei et al (1998) Science 282:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
And Development	<ul style="list-style-type: none"> <li>Nitrogen Metabolism</li> </ul>	287-290; Cutler et al. (1996) Science 273: 1239 Goupil et al (1998) J Exptl Botany 49:1855-62
Water Conservation And Resistance To Drought And Other Related Stresses	<ul style="list-style-type: none"> <li>Stomatal Development And Physiology</li> <li>Stress Response Pathways</li> <li>Inhibition Of Ethylene Production Under Low Water Potential</li> <li>Proline And Other Osmolite Synthesis And Degradation</li> </ul>	Allen et al. (1999) Plant Cell 11: 1785-1798 Li et al. 2000 Science 287: 300-303 Burnett Et Al 2000. J. Exptl Botany 51: 197-205 Raschke (1987) In: Stomatal Function Zeiger et al. Eds., 253-279 Bush And Pages (1998) Plant Mol. Biol. 37: 425-35 Spollen Et Al (2000) Plant Physiol. 122:967-976  Hare et al. (1998) Plant, Cell And Environment 21:535- 553; Hare et al. (1999) J. Exptl. Botany 50:413-434
Programmed cell death	<ul style="list-style-type: none"> <li>Proteases</li> <li>DNA endonucleases</li> <li>Mitochondriae uncoupling proteins</li> </ul>	Kamens et al. (1995) J. Biol. Chem. 270, 15250-15256 Wang et al. (2001) Anticancer Res. 21, 1789- 1794 Drake et al. (1996) Plant

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		<p>Mol. Biol 304, 755-767</p> <p>Mittler and Lam (1995)</p> <p>Plant Cell 7, 1951-1962</p> <p>Mittler and Lam (1995)</p> <p>Plant Physiol. 108, 489-493</p> <p>Thelen and Northcote (1989) Planta 179, 181-195</p> <p>Hanak and Jezek (2001)</p> <p>FEBS Lett. 495, 137-141</p>
	<ul style="list-style-type: none"> <li>• Plasmalemma and Tonoplast Ion Channel Changes</li> <li>• Ca<sup>2+</sup> Accumulation</li> <li>• K<sup>+</sup> Efflux</li> <li>• Activation Of Kinases And Phosphatases</li> </ul>	<p>Macrobbie (1998) Philos Trans R Soc Lond B Biol Sci 353: 1475-88; Li et al (2000) Science 287:300- 303; Barkla et al. (1999)</p> <p>Plant Physiol. 120:811-819</p> <p>Lacombe et al. (2000) Plant Cell 12: 837-51; Wang et al. (1998) Plant Physiol 118:1421-1429; Shi et al. (1999) Plant Cell 11: 2393- 2406</p> <p>Gaymard et al. (1998) Cell 94:647-655</p> <p>Jonak et al. (1996) Proc. Natl. Acad. Sci 93: 11274- 79; Sheen (1998) Proc. Natl. Acad. Sci. 95: 975-80; Allen</p>



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		et al. (1999) Plant Cell 11: 1785-98

Other biological activities that can be modulated by the viability genes, their components and products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

#### IV.A.4. USE OF VIABILITY GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE TRANSCRIPT LEVELS OF OTHER GENES

The expression of many genes is "up regulated" or "down regulated" following herbicide treatment and also in the leaf mutants, because some "viability" genes and their products are integrated into complex networks that regulate transcription of many other genes. Some "viability genes" are therefore useful for modifying the transcription of other genes and hence complex phenotypes, as described above. The data from differential expression experiments can be used to identify a number of types of transcript profiles of "viability genes", including "early responders," and "delayed responders", "early responder repressors" and "delayed repressors".

Profiles of these different types responsive genes are shown in the Table below together with examples of the kinds of associated biological activities. "Up-regulated" profiles are those where the mRNA transcript levels are higher in the herbicide treated plants as compared to the untreated plants. "Down-regulated" profiles represent higher transcript levels in the untreated plant as compared to the herbicide treated plants.

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
Up Regulated Transcripts (Level At 4 Hr $\cong$ 0 Hr) or (Level At 4 Hr > 0 Hr)	<ul style="list-style-type: none"> <li>• Early Responders To:               <ul style="list-style-type: none"> <li>• Gluphosinate</li> <li>• Chlorsulfuron</li> <li>• Glyphosate and/or 2, 4-D</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Suppression of cell, tissue, organ or plant death following:               <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> </ul> </li> <li>• Activation of cell, tissue, organ or plant death following:               <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Transporters</li> <li>• Change In Cell Membrane Structure</li> <li>• Kinases And Phosphatases</li> <li>• Germins, Germin-like proteins, Calcium-binding proteins and H<sub>2</sub>O<sub>2</sub> generating and H<sub>2</sub>O<sub>2</sub> neutralizing proteins.</li> <li>• Transcription Activators</li> <li>• Change In Chromatin Structure And/Or Localized DNA Topology</li> <li>• Annexins, cell wall structural proteins</li> </ul>
Up Regulated Transcripts	Delayed Responders to Gluphosinate,	<ul style="list-style-type: none"> <li>• Suppression of cell, tissue, organ</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
(Level At 4 Hr < 12 Hr)	Chlorsulfuron, Glyphosate and/or 2, 4-D	<p>or plant death following:</p> <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> <li>• Activation of cell, tissue, organ or plant death following: <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Specific Factors (Initiation And Elongation) For Protein Synthesis</li> <li>• Lipid transfer proteins</li> <li>• Myrosinase-binding proteins</li> <li>• Sugar interconverting enzymes</li> <li>• Maintenance Of mRNA Stability</li> <li>• Maintenance Of Protein Stability</li> <li>• Maintenance Of Protein-Protein Interaction</li> <li>• Protein translocation factors</li> <li>• RNA-binding proteins</li> <li>• Centromere and cytoskeleton proteins</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
			<ul style="list-style-type: none"> <li>• Lipases</li> <li>• Zn/Cu transporters</li> <li>• Cell wall structural proteins</li> </ul>
Down-Regulated Transcripts (Level At 0 Hr $\leq$ 4 Hr) or (Level At 0 Hr > 4 Hr)	<p>Early Responder</p> <p>Repressors Of Stress Response State Of Metabolism</p> <p>Genes With Discontinued Expression Or Unstable mRNA In Presence Of Herbicide or Abiotic Stress</p>	<ul style="list-style-type: none"> <li>• Suppression of cell, tissue, organ or plant death following:               <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> </ul> </li> <li>• Activation of cell, tissue, organ or plant death following:               <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> <li>• Zn/Cu transporters</li> <li>• Cell wall structural proteins</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Change In Protein Structure By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>• Change In Chromatin Structure And/Or DNA Topology</li> <li>• H<sub>2</sub>O<sub>2</sub> neutralizing proteins</li> <li>• Neutralizing proteins including SOD and GST</li> </ul>
Down-Regulated Transcripts	Delayed Responder Repressors Of ABA	<ul style="list-style-type: none"> <li>• Suppression of cell, tissue, organ</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES WHOSE TRANSCRIPTS ARE CHANGED	PHYSIOLOGICAL CONSEQUENCES OF MODIFYING GENE PRODUCT LEVELS	EXAMPLES OF BIOCHEMICAL ACTIVITIES WHOSE TRANSCRIPTS ARE CHANGED
(Level At 4 Hr > 12 Hr)	Function State Of Metabolism	<p>or plant death following:</p> <ul style="list-style-type: none"> <li>• Herbicide treatment or under stress</li> <li>• Activation of cell, tissue, organ or plant death following:</li> <li>• Herbicide treatment or under stress</li> </ul>	<ul style="list-style-type: none"> <li>• Kinases And Phosphatases</li> <li>• Stability Of Factors For Protein Synthesis And Degradation</li> <li>• Amino Acid biosynthesis proteins including asparagine synthase</li> </ul>
	Expression Or Unstable mRNA In Presence Of herbicide or Abiotic Stress		<ul style="list-style-type: none"> <li>• Ca-binding proteins</li> <li>• Lipid biosynthesis proteins</li> <li>• Lipases</li> <li>• Zn/Cu transporters</li> <li>• Cell wall structural proteins</li> </ul>

While viability modulating polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development.

IV.A.5. USE OF PROMOTERS OF VIABILITY GENES, GENE COMPONENTS  
AND PRODUCTS

Promoters of viability genes can include those that are induced by (1) destructive chemicals, e.g. herbicides, (2) stress, or (3) death. These promoters can be linked operably to achieve expression of any polynucleotide from any organism. Specific promoters from viability genes can be selected to ensure transcription in the desired tissue or organ. Proteins expressed under the control of such promoters can include those that can induce or accelerate death or those that can protect plant cells organ death. For example, stress tolerance can be increased by using promoters of viability genes to drive transcription of cold tolerance proteins, for example. Alternatively, promoters induced by apoptosis can be utilized to drive transcription of antisense constructs that inhibit cell death.

IV.B. HISTONE DEACETYLASE (AXEL) RESPONSIVE GENES, GENE  
COMPONENTS AND PRODUCTS

The deacetylation of histones is known to play an important role in regulating gene expression at the chromatin level in eukaryotic cells. Histone deacetylation is catalyzed by proteins known as histone deacetylases (HDACs). HDACs are found in multisubunit complexes that are recruited to specific sites on nuclear DNA thereby affecting chromatin architecture and target gene transcription. Mutations in plant HDAC genes cause alterations in vegetative and reproductive growth that result from changes in the expression and activities of HDAC target genes or genes whose expression is governed by HDAC target genes. For example, transcription factor proteins control whole pathways or segments of pathways and proteins also control the activity of signal transduction pathways. Therefore, manipulation of these types of protein levels is especially useful for altering phenotypes and biochemical activities.

Manipulation of one or more HDAC gene activities are useful to modulate the biological activities and/or phenotypes listed below. HDAC genes and gene products can act alone or in combination. Useful combinations include HDAC genes and/or gene products with similar biological activities, or members of the same, co-regulated or functionally related biochemical pathways. Such HDAC genes and gene products can function to either increase or dampen these phenotypes or activities.

Examples of genes whose expression is affected by alterations in HDAC activity are shown in the Reference and Sequence Tables. These genes and/or gene products are responsible for effects on traits such as inflorescence branching and seed production. They were discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA products are affected by a decrease in HDAC gene activity. These experiments made use of an Arabidopsis mutant having severely reduced mRNA levels for the histone deacetylase gene AtHDAC1.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by simultaneously hybridizing two differentially labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC

chips containing 10,000 non-redundant ESTs, selected from 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases over the controls were identified, compared to the Ceres full-length cDNA and genomic sequence databanks, and identical Ceres clones identified. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which are HDAC genes. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Axel (relating to SMD 6654, SMD 6655)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Histone Deacetylase genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Histone Deacetylase Genes Identified By Cluster Analyses Of Differential Expression

##### Histone Deacetylase Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Histone Deacetylase genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Axel (relating to SMD 6654, SMD 6655) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Histone Deacetylase Genes Identified By Correlation To Genes That Cause Physiological Consequences



Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Histone Deacetylase genes. A group in the MA\_clust is considered a Histone Deacetylase pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Histone Deacetylase Genes Identified By Amino Acid Sequence Similarity

Histone Deacetylase genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Histone Deacetylase genes. Groups of Histone Deacetylase genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Histone Deacetylase pathway or network is a group of proteins that also exhibits Histone Deacetylase functions/utilities.

#### IV.B.1.USE OF HDAC GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE PHENOTYPES

HDAC genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth Rate
- Whole Plant, Including Height, Flowering Time, Etc.
- Seedling
- Organ
  - Stem
  - Leaves
  - Roots
  - Flowers
  - Fruits
  - Seeds
  - Yield

- Size, Weight
- Seed Development
- Embryo
- Germination
- Cell Differentiation

5

To improve any of the phenotype(s) above, activities of one or more of the HDAC genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Wu et al. (2000,, Plant J 22: 19-27), Hu et al. (2000, J Biol Chem 275: 15254-64), Johnson and Turner (1999, Semin Cell Dev Biol 10: 179-88), Koyama et al. (2000, Blood 96: 1490-5), Wu et al. (2000, Plant J 22: 19-27), Li (1999, Nature Genetics 23: 5-6), Adams et al. (2000, Development 127: 2493-2502) and Lechner et al. (2000, Biochemistry 39: 1683-92).

#### IV.B.2.USE OF HDAC DEVELOPMENT GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the HDAC genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation And Development	<ul style="list-style-type: none"> <li>• Cell Differentiation</li> <li>• Cell Cycle Progression</li> </ul>	<p>Koyama et al. (2000) Blood 96: 1490-5</p> <p>Hu et al. (2000) J Biol Chem 275: 15254-64</p>
Metabolism	<ul style="list-style-type: none"> <li>• Chromatin Structure</li> </ul>	Hu et al. (2000) J Biol Chem 275: 15254-64

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	<ul style="list-style-type: none"> <li>Gene Transcription And Chromatin Assembly</li> </ul>	Johnson and Turner (1999) Semin Cell Dev Biol 10: 179-88
Reproduction And Seed Development	<ul style="list-style-type: none"> <li>Seed Development</li> <li>Seed Germination</li> <li>Independent Embryo Fertilization</li> <li>Fertilization Independent Seed Development</li> <li>Megagametogenesis</li> </ul>	Wu et al. (2000) Plant J 22:19-27  Lechner et al. (2000) Biochemistry 39: 1683-92 Ohad et al. (1996) PNAS USA 93: 5319-24  Chaudhury et al. (1997) PNAS USA 94: 4222-28  Christensen et al. (1997) Sex Plant Reproduc 10: 49-64

Other biological activities that can be modulated by the HDAC genes and gene products are listed in the REFERENCE Table. Assays for detecting such biological activities are described in the Protein Domain table.

5 Profiles of these different HDAC genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts	Responders To HDAc Activity	<ul style="list-style-type: none"> <li>• Gene Repression Activity</li> <li>• Cell Cycle Progression</li> <li>• Chromatin Condensation</li> <li>• Synthesis Of Metabolites And/Or Proteins</li> <li>• Modulation Of Transduction Pathways</li> <li>• Specific Gene Transcription Initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic enzymes</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> </ul>
Down-Regulated Transcripts	<p>Responder To Hdac Inhibitors</p> <p>Genes With Discontinued Expression Or UnsTable Mrna In Presence Of Hdac</p>	<ul style="list-style-type: none"> <li>• Negative Regulation Of Acetylation Pathways</li> <li>• Changes In Pathways And Processes Operating In Cells</li> <li>• Changes In</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> <li>• Change in chromatin structure</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		Metabolism	and/or DNA topology <ul style="list-style-type: none"> <li>• Stability of factors for protein synthesis and degradation</li> <li>• Metabolic enzymes</li> </ul>

#### Use of Promoters of Histone Deacetylase Responsive Genes

Promoters of Histone Deacetylase responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Histone Deacetylase responsive genes where the desired sequence is operably linked to a promoter of a Histone Deacetylase responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V. STRESS RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

V.A. COLD RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

5 The ability to endure low temperatures and freezing is a major determinant of the geographical distribution and productivity of agricultural crops. Even in areas considered suitable for the cultivation of a given species or cultivar, can give rise to yield decreases and crop failures as a result of aberrant, freezing temperatures. Even modest increases (1-2°C) in the freezing tolerance of certain crop species would have a dramatic impact on agricultural productivity in some areas. The development of genotypes with increased freezing tolerance would provide a more reliable means to minimize crop losses and diminish the use of energy-costly practices to modify the microclimate.

10 Sudden cold temperatures result in modulation of many genes and gene products, including promoters. Examples of such cold responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix tables, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to cold treatment.

20 Manipulation of one or more cold responsive gene activities are useful to modulate the biological activities and/or phenotypes listed below. Cold responsive genes and gene products can act alone or in combination. Useful combinations include cold responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of the levels of such proteins is especially useful for altering phenotypes and biochemical activities of plants. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108578, 108579, 108533, 108534 ). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript

levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Cold genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Cold Genes Identified By Cluster Analyses Of Differential Expression

##### Cold Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Cold genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108578, 108579, 108533, 108534 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Cold Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Cold genes. A group in the MA\_clust is considered a Cold pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Cold Genes Identified By Amino Acid Sequence Similarity

Cold genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Cold genes. Groups of Cold genes are identified in the Protein Group table. In this table, any protein group that

comprises a peptide ID that corresponds to a cDNA ID member of a Cold pathway or network is a group of proteins that also exhibits Cold functions/utilities.

Such cold responsive genes and their products can function to either increase or dampen the phenotypes and activities below either in response to cold treatment or in the absence of cold temperature fluctuations.

Further, promoters of cold responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by ABA or any of the following phenotypes or biological activities below.

#### V.A.1. USE OF COLD-RESPONSIVE GENES TO MODULATE PHENOTYPES

Cold responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Cold Tolerance, below 7°C, for example
- Cells
- Organelles
- Proteins
- Dehydration Resistance
- Growth rate
- Whole Plant, including height, bolting time, etc.
- Organs
  - Roots
  - Lateral Roots
  - Leaves
  - Stems
  - Flowers
  - Fruit
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation;
- Number, Size, and/or Weight
  - Flowers;



- Seeds;
- Branches;
- Leaves;
- Seed Yield
- Number, Size, Weight, Harvest Index, Water Content
- Fruit Yield
- Number, Size, Weight, Harvest Index, Water Content

To regulate any of the phenotype(s) above, activities of one or more of the cold responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance to Steponokus et al. (1993) Biochimica et Biophysica Acta 1145: 93-104; Quinn (1988) Symp Soc. Exp. Biol. 42: 237-258; Bectold and Pelletier (1998) Methods Mol. Biol. 82: 259-266; Kasuga et al. (1999) Nature Biotechnology 17: 287-291; Guy et al. (1998) Cryobiology 36: 301-314; or Liu et al. (1998) Plant Cell 10: 1391-1406.

#### V.A.2. USE OF COLD-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the cold responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations above and those included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cold Tolerance	Viability Of Plant Protoplasts At Low Temperatures.	Steponkus (1998) PNAS USA 95: 14570-14575
	Viability Of Yeast At Low Temperatures.	Schirmer et al. (1994) Plant Cell 6: 1899-1909
	Complementation Of Yeast Tsp Mutant	Zentella et al. (1999) Plant Physiology, 119: 1473-1482
	Viability Of E.Coli At Low Temperatures.	Yeh et.al. (1997) PNAS 94: 10967-10972
	Induction Of Cold Shock Response Genes	Pearce (1999) Plant Growth Regulation 29: 47-76.
Lipid Composition	Altered Composition Of Membrane Fatty Acids	Sayanova et al. (1999) Journal of Experimental Botany 50: 1647-1652 Sayanova (1997) PNAS USA 94: 4211-4216
	Alteration Of Lipoyxygenase Enzyme Accumulation And Activity	Porta et al. (1999) Plant and Cell Physiology 40: 850-858.
Protein Composition	<ul style="list-style-type: none"> <li>- Protein Denaturation</li> <li>- Protein Hydrophilicity</li> </ul>	Wisniewski et al.(1999) Physiologia Plantarum 105: 600-608 Steponkus (1998) PNAS USA 95: 14570-14575

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Modulation of Transcription Induced by Low Temperatures	<ul style="list-style-type: none"> <li>- Induced Transcription Factors And Other Dna Binding Proteins</li> <li>- Transcription Of Specific Genes</li> </ul>	<p>Current Protocols in Molecular Biology / edited by Frederick M. Ausubel .. [et al.]. New York : Published by Greene Pub. Associates and Wiley-Interscience : J. Wiley, c1987.</p> <p>Steponkus (1998) PNAS USA 95: 14570-14575</p> <p>Kadyrzhanova et al., Plant Mol Biol (1998) 36(6): 885-895; and</p> <p>Pearce et al., Plant Physiol (1998) 117(3): 787-795</p>
Signal Transduction	Plasma Membrane Proteins	<p>Goodwin et al., Plant Mol Biol (1996) 31(4) 777-781; and</p> <p>Koike et al., Plant Cell Physiol (1997) 38(6): 707-716</p>
Oxygen Scavengers	<ul style="list-style-type: none"> <li>- Glutathione</li> <li>- Accumulation Active O<sub>2</sub> and H<sub>2</sub>O<sub>2</sub> Scavengers</li> </ul>	<p>Kocsy et al., Planta (2000) 210(2): 295-301</p> <p>Tao et al., Cryobiology (1998) 37(1):38-45</p>
Dehydration	<ul style="list-style-type: none"> <li>- Dehydrin</li> <li>- Transcription of mRNA</li> </ul>	<p>Ismail et al., Plant Physiol (1999) 120(1):237-244</p> <p>Kaye et al., Plant Physiol (1998) 116(4): 1367-1377</p>
Metabolism	Soluble Sugars and/or Proline	Wanner et al., (1999) Plant Physiol 120(2): 391-400

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
RNA/DNA Chaperone	Stabilization of RNA/DNA through RNA binding and modulation of RNA translation through RNA binding and or unwinding.	Jiang, Weining et al.,(1997) Journal of Biological Chemistry, 272: 196-202. Fukunaga et al., (1999) Journal of Plant Research, 112: 263-272.
Protein Chaperone	Stabilize protein structure and facilitate protein folding	Forreiter and Nover (1998) Journal of Biosciences 23: 287-302

Other biological activities that can be modulated by the cold responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Cold responsive genes are characteristically differentially expressed in response to fluctuating cold temperature levels, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various cold responsive genes in the aerial parts of seedlings at 1 and 6 hours at 4°C in the dark as compared to aerial parts of seedlings covered with aluminium foil, and grown at 20°C in the growth chamber.

The data from this time course can be used to identify a number of types of cold responsive genes and gene products, including "early responders" and "delayed responders". Profiles of these different cold responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Upregulated Genes (Level At 1 h $\approx$ 6 h) or (Level At 1 h > 6 h)	Early Responders To Cold	<ul style="list-style-type: none"> <li>- Perception Of Cold</li> <li>- Induction Of Cold Response</li> <li>- Transduction Pathways</li> <li>- Initiating Specific Gene Transcription</li> <li>- Osmotic Adjustment</li> <li>- Alteration Of Lipid Composition.</li> <li>- Ice Nucleation Inhibition</li> <li>- Mitigation Of Dehydration By Sequestering Water</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Kinases And Phosphatases</li> <li>- Amino Acid Sugar And Metabolite Transporters</li> <li>- Carbohydrate Catabolic And Anabolic Enzymes.</li> <li>- Lipid Biosynthesis Enzymes</li> <li>- Lipid Modification Enzymes, Example Desaturases</li> <li>- Ice Crystal Binding Proteins</li> <li>- Hydrophilic Proteins</li> </ul>

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
	Stress Response	<ul style="list-style-type: none"> <li>- Repression Of General Biochemical Pathways To Optimize Cold Response Pathways.</li> <li>-Stabilization Of Protein /Enzyme Activity At Low Temperature</li> <li>-Protection Against Oxidative Stress</li> <li>-Anaerobic Metabolism</li> </ul>	<ul style="list-style-type: none"> <li>-Transcription Factors</li> <li>- Kinases And Phosphatases</li> <li>- Protein Stability Factors</li> <li>- mRNA Stability Factors</li> <li>- mRNA Translation Factors</li> <li>- Protein Turnover Factors</li> <li>Oxygen Radical Scavengers, Example- Peroxidases</li> <li>- Energy Generation Enzymes EtOH Detoxification</li> </ul>

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Upregulated Genes (Level At 1h < 6 h)	Delayed Responders To Cold Stress - Cold Acclimation Genes	- Respiration, Photosynthesis And Protein Synthesis - Carbohydrate And Amino Acid Solute Accumulation - Increased Fatty Acid Desaturation To Increase Lipid Membrane Stability - Increased Accumulation Or Activity Of Oxidative Stress Protection Proteins - Stabilization Of Protein /Enzyme Activity At Low Temperature - Protection Against Oxidative Stress - Extracellular Matrix Modification	-Transcription Factors - Kinases And Phosphatases - Protein Stability Factors - mRNA Stability Factors - mRNA Translation Factors - Protein Turnover Factors - Oxygen Radical Scavengers, Peroxidase - Metabolic Enzymes

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
	Stress Response Genes	<ul style="list-style-type: none"> <li>- Stabilization Of Protein /Enzyme Activity At Low Temperature</li> <li>- Protection Against Oxidative Stress</li> <li>- Anaerobic Metabolism</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Kinases And Phosphatases</li> <li>- Protein Stability Factors</li> <li>- mRNA Stability Factors</li> <li>- mRNA Translation Factors</li> <li>- Protein Turnover Factors</li> <li>- Oxygen Radical Scavengers, Example- Peroxidase</li> <li>- Energy Generation Enzymes, Etoh</li> <li>- Detoxification</li> </ul>



GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Downregulated (Level At 1 h $\approx$ 6 h) (Level At 6 h > 1 h)	<ul style="list-style-type: none"> <li>- Early Responder Repressors Of Cold Stress Metabolism</li> <li>- Genes With Discontinued Expression Or UnsTable mRNA In Cold</li> </ul>	<ul style="list-style-type: none"> <li>- Negative Regulation Of Cold Signal Transduction Pathways Released</li> <li>- Negative Regulation Of Cold Induced Transcription Reduced</li> <li>- Reduction In Gene Expression In Pathways Not Required Under Cold Conditions</li> <li>- Induced mRNA Turnover</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Kinases And Phosphatases</li> <li>- Protein Stability Factors</li> <li>- mRNA Stability Factors</li> <li>- mRNA Translation Factors</li> <li>- Protein Turnover Factors</li> <li>- Cold Repressed Metabolic Pathway Proteins</li> <li>- Factors Coordinating And Controlling Central C and N Metabolism</li> <li>- Storage Proteins</li> </ul>

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Down-Regulated Transcripts (Level At 1 h > 6 h)	<ul style="list-style-type: none"> <li>- Delayed Responder</li> <li>- Repressors Of Cold Stress Metabolism</li> <li>- Genes With Discontinued Expression Or UnsTable mRNA In Cold</li> </ul>	<ul style="list-style-type: none"> <li>- Maintenance Of Cold Induced State Of Metabolism</li> <li>- Reduction In Gene Expression For Pathways Not Required Under Cold Conditions</li> <li>- Induced mRNA Turnover</li> </ul>	<ul style="list-style-type: none"> <li>- Transcription Factors</li> <li>- Kinases And Phosphatases</li> <li>- Protein Stability Factors</li> <li>- mRNA Stability Factors</li> <li>- mRNA Translation Factors</li> <li>- Protein Turnover Factors</li> <li>- Cold Repressed Metabolic Pathway Proteins</li> <li>- Factors Coordinating And Controlling Central C and N Metabolism</li> <li>- Storage Proteins</li> </ul>

Further, any desired sequence can be transcribed in similar temporal, tissue, or enviromentally specific patterns as the cold responsive genes when the desired sequence is operably linked to a promoter of a cold responsive gene.

V.B. HEAT RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

The ability to endure high temperatures is a major determinant of the geographical distribution and productivity of agricultural crops. Decreases in yield and crop failure frequently occur as a result of aberrant, hot conditions even in areas considered suitable for the cultivation of a given species or cultivar. Only modest increases in the heat tolerance of crop species would have a dramatic impact on agricultural productivity. The development of genotypes with increased heat tolerance would provide a more reliable means to minimize crop losses and diminish the use of energy-costly practices to modify the microclimate.

Changes in temperature in the surrounding environment or in a plant microclimate results in modulation of many genes and gene products. Examples of such heat stress responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to high temperatures.

While heat stress responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different heat stress responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a heat stress responsive polynucleotide and/or gene product with other environmentally responsive polynucleotide is also useful because of the interactions that exist between stress pathways, pathogen stimulated pathways, hormone regulated pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles, but which participate in common or overlapping pathways. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108576, 108577, 108522, 108523). For transcripts that had higher levels in the samples than the control,

a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Heat genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Heat Genes Identified By Cluster Analyses Of Differential Expression

##### Heat Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Heat genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108576, 108577, 108522, 108523 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Heat Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Heat genes. A group in the MA\_clust is considered a Heat pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Heat Genes Identified By Amino Acid Sequence Similarity

Heat genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Heat genes. Groups of Heat genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Heat pathway or network is a group of proteins that also exhibits Heat functions/utilities.

Such heat stress responsive genes and gene products can function either to increase or dampen the above phenotypes or activities either in response to changes in temperature or in the absence of temperature fluctuations.

Further, promoters of heat responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by heat or any of the following phenotypes or biological activities below.

V.B.1. USE OF HEAT STRESS RESPONSIVE GENES TO MODULATE  
PHENOTYPES

Heat stress responsive genes and gene products can be used to alter or modulate one or more of the following phenotypes:

- Heat Tolerance, above 20°C, 23°C, 27°C, 30°C, 33°C, 37°C, 40°C or 42°C
- Of Cells
- Of Organelles
- Of Proteins
- Dehydration Resistance
- Of Cells
- Of Organelles
- Growth Rate
- Whole Plant, including height, bolting time, etc.
- Organs
  - Roots
  - Lateral Roots
  - Leaves
  - Stems
  - Flowers
  - Fruit
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation;
- Number, Size, and Weight of
  - Flowers;

- Seeds;
- Branches;
- Leaves;
- Seed Yield
- Number, Size, Weight, Harvest Index
- Fruit Yield
- Number, Size, Weight, Harvest Index
- Stress Responses
  - Mediation of response to desiccation, drought, salt, disease, wounding, cold and other stresses.
- Reproduction

To regulate any of the phenotype(s) above, activity of one or more of the heat stress responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Queitsch et al. (2000, The Plant Cell 12: 479-92).

#### V.B.2. USE OF HEAT STRESS RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the heat stress responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below.

Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
Cell Growth and Differentiation	-Regulation And Molecular Chaperones	Wisniewski et al. (1999) Physiolgia

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
	<p>-Maintenance Of Native Conformation (Cytosolic Proteins)</p> <p>-Reactivation Of Aggregation And Protein Folding</p> <p>-Autoregulation Of Heat Shock Response</p> <p>-Regulation Of Translational Efficiency</p> <p>-Regulation Of Kinase Activity</p> <p>-Regulation Of Calcium Mediated Signal Transduction</p>	<p>Plantarum 105: 600-608</p> <p>Queitsch et al. (2000) The Plant Cell 12: 479-92</p> <p>Lee and Vierling (2000) Plant Physiol. 122: 189-197</p> <p>Schwechheimer (1998) Plant Mol Biol 36: 195-204</p> <p>Shi et al. (1998) Genes and Development 12: 654-66</p> <p>Wells et al. (1998) Genes and Development 12: 3236-51</p> <p>Lis et al. (2000) Genes and Development 14: 792-803</p> <p>-Malho, R.(1999) Plant Biology 1: 487-494.</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
		<p>-Sheen, Jen.(1996) Science 274: 1900-1902.</p> <p>- Farmer, P. et al., (1999.) Biochimica et Biophysica Acta 1434: 6-17.</p>
Gene regulation	<ul style="list-style-type: none"> <li>• Transcriptional Regulation Of Heat Induced Proteins Through DNA Binding Proteins.</li> <li>• Transcriptional Regulation Of Heat Induced Proteins Through Protein-Protein Interactions Between DNA Binding Proteins And Coactivators.</li> </ul>	<p>-Current Protocols in Molecular Biology / edited by Frederick M. Ausubel .. [et al.]. New York : Published by Greene Pub. Associates and Wiley-Interscience : J. Wiley, c1987.</p> <p>-Steponkus (1998) PNAS USA 95: 14570-14575</p> <p>- Gubler et al. (1999) Plant Journal 17: 1-9</p> <p>- Glenn et al. (1999) Journal of Biological Chemistry, 274: 36159-36167</p>



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
	<ul style="list-style-type: none"> <li>• Transcriptional Regulation Of Heat Induced Proteins Through Protein Phosphorylation And Dephosphorylation</li> <li>• Transcriptional Regulation Of Thermal Stress Induced Genes By Protein-Protein Interactions.</li> <li>• Translational Regulation Of Thermal Stress Induced Messenger Rnas.</li> <li>• Transcriptional Regulation Of Heat Induced Genes Through Chromatin Remodeling.</li> </ul>	<p>- Zhou et al., (1997) EMBO Journal 16:3207-3218.</p> <p>- Sessa et al., (2000) EMBO Journal 19: 2257-2269.</p> <p>- Burnett et al., (2000) Journal of Experimental Botany. 51: 197- 205.</p> <p>- Osterlund et al., (2000) Nature 405: 462-466.</p> <p>- Gross and Watson (1998) Canadian Journal of Microbiology, 44:341-350</p> <p>- Luo, R. X., Dean, D.C. (1999) Journal of the</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
		National Cancer Institute 91: 1288- 1294.  -Chromatin protocols (1999) edited by Peter B. Becker. Totowa, N.J. : Humana Press.
Cell Structure	<ul style="list-style-type: none"> <li>Thermal Stress Protection By Plasma Membrane Anchored Or Secreted And/Or Cell Wall Associated Proteins.</li> </ul>	- Goodwin et al. (1996) Plant Mol Biol 31(4) 777-781; and Koike et al. (1997) Plant Cell Physiol 38(6): 707-716
Signal Transduction	<ul style="list-style-type: none"> <li>Regulation Of Thermal Stress Pathways And Protein Activity By Protein Kinase And Protein Phosphatase Mediated Phosphorylation And Dephosphorylation Respectively.</li> </ul>	- Jonak (1996) Proceedings of the National Academy of Sciences of the United States of America, 93: 11274-11279. - Monroy et al., (1998) Analytical Biochemistry 265: 183-185.
Photosynthesis	<ul style="list-style-type: none"> <li>Regulation Of Photoprotection And Repair Of Photosystem II</li> </ul>	Schroda et al. (1999) The Plant Cell 11: 1165-178 Oh and Lee (1996) J Plant Biol. 39: 301-07
Stress Response	<ul style="list-style-type: none"> <li>Regulation Of Cytosol</li> </ul>	Dat et al. (1998) Plant

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
	<p>Peroxide Levels</p> <ul style="list-style-type: none"> <li>Regulation Of Heat Shock Factor Binding</li> <li>Regulation Of Protein Stability During Thermal Stress</li> <li>Nucleocytoplasmic Export Of Heat Shock Protein Mrnas</li> <li>Regulation/Reconfiguration Of Cell Architecture</li> <li>Regulation Of Pathways For Reactivation Of "Damaged" And/Or Denatured Proteins</li> </ul>	<p>Physiol 116: 1351-1357</p> <p>Kurek et al. (1999) Plant Physiol 119: 693-703</p> <p>Storozhenko et al. (1998) Plant Physiol 118: 1005-14</p> <p>Soto et al. (1999) Plant Physiol 120: 521-28</p> <p>Yeh et al. (1997) PNAS 94: 10967-10972</p> <p>Winkler et al. (1998) Plant Physiol 118: 743-50</p> <p>Saavedra et al. (1997) Genes and Development 11: 2845-2856</p> <p>Parsell and Lindquist (1993). Ann. Rev. Genet. 27: 437-496.</p> <p>Parsell and Lindquist (1993). Ann. Rev. Genet. 27: 437-496.</p> <p>Georgopoulos and Welch (1993). Ann Rev. Cell Biol. 9:601-634.</p>

003556 082401

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAY
	<ul style="list-style-type: none"> <li>• Regulation Of Protein Degradation During Thermal Stress.</li> <li>• Regulation Of Osmotic Potential During Thermal Stress.</li> <li>• Regulation Of Universal Stress Protein Homologue Activity By Phosphorylation And Dephosphorylation.</li> <li>• Regulation Of Dehydrin, LEA-Like And Other Heat Stable Protein Accumulation</li> </ul>	<p>- Vierstra, Richard D. (1996) Plant Molecular Biology,32:275-302.</p> <p>- Vierstra, Richard D.; Callis, Judy. (1999) Plant Molecular Biology, 41:435-442.</p> <p>- Liu, J. et al., (1998)Plant Science 134:11-20.</p> <p>- Freestone, P. 1997et al., Journal of Molecular Biology, v. 274: 318-324.</p> <p>- Robertson,A.J. (1994) Plant Physiology 105: 181-190.</p>

Other biological activities that can be modulated by the heat stress responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Heat stress responsive genes are characteristically differentially transcribed in response to fluctuating temperatures, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various heat stress responsive genes in aerial tissues at 1 and 6 hours after plants were placed at 42°C as compared to aerial tissues kept at 20°C growth chamber temperature.

The data from this time course can be used to identify a number of types of heat stress responsive genes and gene products, including "early responders to heat stress," "delayed responders to heat stress," "early responder repressors," and "delayed repressor responders." Profiles of these different heat stress responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES / GENE PRODUCTS
Up Regulated Transcripts (Level At 1h ≈ 6h) Or (Level At 1h > 6h)	Early Responders To Heat Stress	<ul style="list-style-type: none"> <li>Heat Stress Perception</li> <li>Modulation Of Heat Stress Response Transduction Pathways</li> <li>Specific Gene Transcription Initiation</li> <li>Conditional Shift In Preferential Translation Of Transcripts</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>Transporters</li> <li>Changes In Cell Membrane Structure</li> <li>Kinases And Phosphatases</li> <li>Transcription Activators</li> <li>Changes In Chromatin Structure And/Or Localized Dna Topology</li> </ul>

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES / GENE PRODUCTS
		<ul style="list-style-type: none"> <li>Changes In Cell Architecture To Optimize Cell Adaptation To Heat Stress</li> </ul>	<ul style="list-style-type: none"> <li>Modification Of Pre-Existing Translation Factors By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>Synthesis Of New Translation Factors</li> <li>Stability Of Mediators Of Protein-Protein Interaction</li> </ul> <p>Heat Shock Proteins Changes In Organelle Structures, Membranes And Energy-Related Activities</p> <p>Proteins To Catalyse Metabolic Turnover</p>
Up Regulated	"Delayed"		<ul style="list-style-type: none"> <li>Transcription</li> </ul>

# Chengdu

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES / GENE PRODUCTS
	Discontinued Expression Or UnsTable mRNA In Presence Of Heat Stress	Pathways And Processes Operating In Cells • Reorientation Of Metabolism	(Kinases) Or Dephosphorylation (Phosphatases) • Change In Chromatin Structure And/Or Dna Topology
Down-Regulated Transcripts (Level At 1hr > 6hr)	Delayed Repressors Of "Normal" State Of Metabolism  Genes With Discontinued Expression Or UnsTable mRNA In Presence Of Heat Stress	• Maintenance Of Heat Stress Response • Maintenance Of Pathways Released From Repression • Changes In Pathways And Processes Operating In Cells • Reorientation Of Metabolism	• Transcription Factors And Activators • Kinases And Phosphatases • Stability Of Factors For Protein Translation

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the heat responsive genes when the desired sequence is operably linked to a promoter of a heat responsive gene.



V.C. DROUGHT RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

The ability to endure drought conditions is a major determinant of the geographical distribution and productivity of agricultural crops. Decreases in yield and crop failure frequently occur as a result of aberrant, drought conditions even in areas considered suitable for the cultivation of a given species or cultivar. Only modest increases in the drought tolerance of crop species would have a dramatic impact on agricultural productivity. The development of genotypes with increased drought tolerance would provide a more reliable means to minimize crop losses and diminish the use of energy-costly practices to modify the microclimate.

Drought conditions in the surrounding environment or within a plant, results in modulation of many genes and gene products. Examples of such drought responsive genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to availability of water.

While drought responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different drought responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways, or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of the levels of such proteins is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a drought responsive polynucleotide and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in a common pathway. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108572, 108573, 108502, 108503, 108504, 108556, 108482, 108483, 108473, 108474, 108477). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript

levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Drought genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Drought Genes Identified By Cluster Analyses Of Differential Expression

##### Drought Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Drought genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108572, 108573, 108502, 108503, 108504, 108556, 108482, 108483, 108473, 108474, 108477 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Drought Genes Identified By Correlation To Genes That Cause

##### Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Drought genes. A group in the MA\_clust is considered a Drought pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Drought Genes Identified By Amino Acid Sequence Similarity

Drought genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Drought genes. Groups of Drought genes are identified in the Protein Group table. In this table, any protein group that

comprises a peptide ID that corresponds to a cDNA ID member of a Drought pathway or network is a group of proteins that also exhibits Drought functions/utilities.

Such drought responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to drought conditions or in the absence of drought conditions. Further, promoters of drought responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by drought or any of the following phenotypes or biological activities below.

More specifically, drought responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
- Roots
- Stems
- Buds
- Leaves
- Development
- Cell Growth
- Leaves
- Fruit Development
- Seed Development
- Senescence
- Stress Responses
- Mediates response to desiccation, drought, salt and cold

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the drought responsive genes when the desired sequence is operably linked to a promoter of a drought responsive gene.

To produce the desired phenotype(s) above, one or more of the drought response genes or gene products can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-

266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Ruzin (1999, In: Plant Microtechnique and Microscopy, Oxford University Press, London) and Khanna-Chopra et al. (1999, BBRC 255:324-7).

5

Alternatively, the activities of one or more of the drought responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
Cell Growth and Differentiation	Preservation of Leaf Sub-Cellular Structures Including Photosynthetic Apparatus	Jagtap et al. (1998) J Exptl Botany 49:1715-1721
	Preservation of Cell Membrane Structures	Munne-Bosch and Alegre (2000) Planta 210: 925-31
	Regulation of Stomatal development and Physiology	Menke et al. (2000) Plant Physiol. 122:677-686.
	Regulation of Factors Involved in the Drought-adapted change in cell ultrastructure	Harrak et al. (1999) Plant Physiol. 121:557-564.
Physiology	Modulation of Transpiration	Allen et al. (1999) Plant Cell 11: 1785-98 Li et al. (2000) Science 287: 300-303 Burnett et al. (2000) J Exptl Bot 51: 197-205

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
	<p>Modulation of Photosynthesis</p> <p>Regulation of Epicuticular Wax Biosynthesis</p> <p>Regulation of Carotenoid Biosynthesis</p>	<p>Raschke (1987) In: Stomatal function, Zeiger et al., Eds, 253-79</p> <p>Sung and Krieg (1979) Plant Physiol 64: 852-56</p> <p>Rhee et al. (1998) Plant Physiol 116: 901-11</p> <p>Alegre (2000) Planta 210: 925-31</p> <p>Loggini et al (2000) Plant Physiol 119:1091</p>
Stress Response	<p>Modulation of Leaf Rolling to minimize water loss</p> <p>Modulation of Osmolite Synthesis</p> <p>Regulation of gene transcriptional activity specific to</p>	<p>Taiz and Zeiger (1991) In: Plant Physiology, Benjamin/Cummings Publishing Co., Redwood City, pp 346-70</p> <p>Hare et al. (1998) Plant, Cell and Environment 21: 535-553</p> <p>Huan et al. (2000) Plant Physiol 122: 747-756</p> <p>Hare et al. (1999) J. Exptl. Botany 333:413-434.</p>

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
	the establishment of drought tolerance	
	Regulation of protein degradation and reactivation during drought stress condition	Lee and Vierling (2000) Plant Physiol. 122: 189-197
	Modulation/reconfiguration of translation machineries ("recycling" mechanisms) adapTable to drought condition	Lis et al. (2000) Genes and Development 14: 792-803
Signal Transduction	Regulation of Ion Sequestration	Bush and Jones (1987) Cell Calcium 8: 455-72
	Regulation of Nuclear Targeted Protein Transport	Ferringno and Silver (1999) Methods in Cell Biology 58: 107-22
	Regulation of Cytoplasmic Ca <sup>2+</sup>	Shi et al. (1999) Plant Cell 11: 2393-2406
	Regulation of Kinase Synthesis and Activity	Li et al. (2000) Science 287- 300-03
	Modulation of Molecular Chaperone Activity	Mayhew et al (1996) Nature 379: 420-26

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
		Kimura et al. (1995) Science 268:1362-1365.

Other biological activities that can be modulated by the drought responsive genes and gene products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Protein Domain table.

Drought responsive genes are characteristically differentially transcribed in response to drought conditions, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various drought responsive genes at 1 and 6 hours after aerial tissues were isolated and left uncovered at room temperature on 3 MM paper, as compared to isolated aerial tissues placed on 3 MM paper wetted with Hoagland's solution.

The data from this time course can be used to identify a number of types of drought responsive genes and gene products, including "early responders," and "delayed responders." Profiles of these different drought responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Up regulated transcripts (level at 1 hr $\approx$ 6 hr) (level at 1 hr > 6 hr)	Early responders to drought	Drought perception leading to the establishment of tolerance to drought  Modulation of drought response transduction	Transcription factors Transporters  Change in cell membrane structure

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
		pathways	Kinases and phosphatases
		Specific gene transcription initiation	Transcription activators Change in chromatin structure and/or localized DNA topology
		Conditional shift in preferential translation of transcripts	Modification of pre- existing translation factors by phosphorylation (kinases) or dephosphorylation (phosphatases) Synthesis of new translation factors
		Changes in cell architecture to optimize cell adaptation to heat stress	Stability of mediators of protein-protein interaction
		Changes in cell division cycle	Synthesis and/or stability of factors regulating cell division
Up regulated transcripts (level at 1 hr < 6 hr)	Maintenance of drought response	Maintenance of response to drought and maintenance of	Transcription factors Specific factors (initiation and elongation) for protein



GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
	"Delayed" responders	<p>drought-tolerance mechanisms</p> <p>Maintenance of mechanisms effective for ions sequestration, osmolite biosynthesis, nuclear protein transport, regulation of cytoplasmic <math>Ca^{+2}</math>, and regulation of proteins effective for maintaining protein stability and conformation</p> <p>Maintenance of cellular structures</p>	<p>synthesis</p> <p>RNA-binding proteins effective for mRNA stability</p> <p>Change in chromatin structure and/or DNA topology</p> <p>Stability of mediators of protein-protein interaction</p> <p>Stability of factors to effectively utilize pre-existing translation machinery ("recycling" mechanisms) under drought condition</p> <p>Stability of mediators of protein-protein interaction</p>
Down-regulated transcripts	Early responder repressors of "normal"	Negative regulation of drought response	Transcription factors and activators

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
(level at 1 hr $\approx$ 6 hr) (level at 6 hr > 1 hr)	state of metabolism  Genes with discontinued expression or unstable mRNA in presence of water stress	inducible pathways released Changes in biochemical and signal transduction pathways and processes operating in cells	Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases) Change in chromatin structure and/or DNA topology
Down-regulated transcripts (level at 1 hr > 6 hr)	Delayed repressors of "normal" state of metabolism  Genes with discontinued expression or unstable mRNA in presence of water stress	Maintenance of drought response Maintenance of pathways released from repression Changes in pathways and processes operating in cells	Transcription factors and activators Kinases and phosphatases Stability of factors for protein translation

#### Use of Promoters of Drought Responsive Genes

Promoters of Drought responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Drought responsive genes where the desired sequence is operably linked to a promoter of a Drought responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such

promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.D. WOUNDING RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Plants are continuously subjected to various forms of wounding from physical attacks including the damage created by pathogens and pests, wind, and contact with other objects.

Therefore, survival and agricultural yields depend on constraining the damage created by the wounding process and inducing defense mechanisms against future damage.

Plants have evolved complex systems to minimize and/or repair local damage and to minimize subsequent attacks by pathogens or pests or their effects. These involve stimulation of cell division and cell elongation to repair tissues, induction of programmed cell death to isolate the damage caused mechanically and by invading pests and pathogens, and induction of long-range signaling systems to induce protecting molecules, in case of future attack. The genetic and biochemical systems associated with responses to wounding are connected with those associated with other stresses such as pathogen attack and drought.

Wounding results in the modulation of activities of specific genes and, in consequence, of the levels of key proteins and metabolites. These genes, called here wounding responsive genes, are important for minimizing the damage induced by wounding from pests, pathogens and other objects. Examples of such wounding responsive genes, gene components and products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff, and MA\_clust tables. They can be active in all parts of a plant and so where, when and to what extent they are active is crucial for agricultural performance and for the quality, visual and otherwise, of harvested products. They were discovered and characterized from a much larger set of genes by experiments designed to find genes whose products changed in response to wounding.

Manipulation of one or more wounding responsive gene activities are useful to modulate the biological activities and/or phenotypes listed below. Wounding responsive genes and gene products can act alone or in combination with genes induced in other ways. Useful combinations include wounding responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of the levels of such proteins is especially useful for altering phenotypes and biochemical activities of plants.

5 The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108574, 108575, 108524, 108525, and Wounding (relating to SMD 3714, SMD 3715) ). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Wounding genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

10 Wounding Genes Identified By Cluster Analyses Of Differential Expression

15 Wounding Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

20 A pathway or network of Wounding genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108574, 108575, 108524, 108525, and Wounding (relating to SMD 3714, SMD 3715) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

25 Wounding Genes Identified By Correlation To Genes That Cause Physiological Consequences

30 Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Wounding genes. A group in the MA\_clust is considered a Wounding pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Wounding Genes Identified By Amino Acid Sequence Similarity

Wounding genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Wounding genes. Groups of Wounding genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Wounding pathway or network is a group of proteins that also exhibits Wounding functions/utilities.

Such wounding responsive genes and gene products can function either to increase or dampen the phenotypes and activities below, either in response to wounding or in the absence of wounding.

Further, promoters of wounding responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by wounding or any of the following phenotypes or biological activities below.

V.D.1. USE OF WOUNDING-RESPONSIVE GENES TO MODULATE  
PHENOTYPES

Wounding responsive genes and gene products can be used to alter or modulate one or more of the following phenotypes:

- Growth Rate
- Whole Plant
  - Height
  - Width
  - Flowering Time
- Organs
  - Coleoptile Elongation
  - Young Leaves
  - Roots
  - Lateral Roots
  - Tuber Formation
  - Flowers
  - Fruit

- Seeds

- Biomass
- Fresh And Dry Weight During Any Time In Plant Life, Such As At Maturation
- Number Of Flowers
- Number Of Seeds
- Seed Yield
- Number
- Size
- Weight
- Harvest Index
  - Content and Composition, e.g., Amino Acid, Nitrogen, Oil, Protein, And Carbohydrate
- Fruit Yield
- Number
- Size
- Weight
- Harvest Index
- Post Harvest Quality
- Content And Composition, e.g., Amino Acid, Carotenoid, Jasmonate, Protein, And Starch
- Seed and Fruit Development
- Germination Of Dormant And Non-Dormant Seeds
- Seed Viability
- Seed Reserve Mobilization
- Fruit Ripening
- Initiation Of The Reproductive Cycle From A Vegetative State
- Flower Development Time
- Insect Attraction For Fertilization
- Time To Fruit Maturity
- Senescence
- Fruits, Fruit Drop

- Leaves
- Stress And Disease Responses
- Drought
- Heat And Cold
- Wounding By Any Source, Including Wind, Objects, Pests And Pathogens
- Uv And High Light Damage
  - Insect, Fungus, Virus, Worm, Nematode Damage

To regulate any of the phenotype(s) above, activities of one or more of the wounding responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance with Johnson et.al. (1998) Plant Physiol 116:643-649, Reymond et.al. (2000) Plant Cell 12 707-720, or Keith et.al. (1991) Proc. Nat. Acad. Sci.USA 888821 8825.



V.D.2. USE OF WOUNDING-RESPONSIVE GENES TO MODULATE  
BIOCHEMICAL ACTIVITIES

The activities of one or more of the wounding responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations included in the Table below:

PROCESS	BIOLOGICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Plant Tissue Proliferation	Cell Damage Repair; Cell Division	Flanders (1990) J. Cell Biol. 110: 1111-1122
Wound Induced Pathways Providing Defense Against Pests And Pathogens	Synthesis Of Jasmonic And Salicylic Acids And The Pathways Induced By These Signaling Molecules. Induction Of Jasmonic Acid Independent Defense Pathways. Induction Of Lipxygenase, Thionins And Nodulins	Reymond, P and Farmer E.E. Current Opinion in Plant Biology 1998 1:404-411 Creelman, RA and Mullet, J.E. (1997) Ann Rev. Plant Physiol Mol Biol 48: 355-387 Leon et al. 1998 Mol Gen Genet 254: 412-419 Titarentko et al. 1997 Plant Physiol 115: 817-826
	Cell Wall Degradation, Ethylene Formation, Systemic Signaling And Induction Of Defense Related Genes	Rojo, E. et al. 1998. Plant J 13:153-165 Ryan, CA and Pearce, G. 1998. Ann Rev. Cell Dev. Biol 14: 1-17
	Specific Rnase Induction	Reymond, P. et al. 2000 . Plant Cell 12:707-720 Glazebrook, J. 1999. Current Opinion in Plant Biol. 2: 280-

PROCESS	BIOLOGICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		286 O'Donnel P. J., et al. 1996 Science 274: 1914-1917 Rojo et al. 1999. Plant J. 20: 135-142 Merkouropoulus G. et al. 1999 Planta 208: 212-219 Kariu et al. 1998. Bioscience Biotechnology and Biochemistry 62: 1144-1151 Mcoann et al. 1997 PNAS 94: 5473-5477
Other Stress Induced Pathways	Absciscic Acid Formation And Its Signaling Pathway Cold Responsive Genes and Pathways Drought Induced Dehydrins And Pathways	Carrera, E and Prat, S. 1998. Plant J 15: 767-771 Chao et. al. 1999. Plant Physiol 120: 979-992
Modified Lipid Metabolism	Membrane Lipid Synthesis Including Omega-3 Fatty Acid Desaturase Lipases Lipid Transfer Proteins	Martin, M et al. 1999 Europe J. Biochem 262: 283-290
Modified Sugar And Energy Metabolism	Induction Of Glycohydrolases And Glycotransferases, Amylases Induction Of	

PROCESS	BIOLOGICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Modified Protein And Nitrogen Metabolism	Aminotransferases, Arginase, Proteases And Vegetative Storage Proteins, Aromatic Amino Acid Synthesis	
Secondary Metabolite Induction	Aromatic Amino Acid Synthesis And Secondary Metabolites	Keith, B et al. 1991 PNAS 88: 8821-8825

Other biological activities that can be modulated by wound responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various wound responsive genes in the aerial parts of a plant, 1 and 6 hours after the plants were wounded with forceps. The comparison was made with aerial tissues from unwounded plants.

The data from this time course reveal a number of types of wound responsive genes and gene products, including "early responders," and "delayed responders." Profiles of the individual wounding responsive genes are shown in the Table below together with examples of the kinds of associated biological activities that are modulated when the activities of one or more such genes vary in plants.

TRANSCRIPT LEVELS	TYPES OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up Regulated Transcripts (Level At 1h $\approx$ 6h) Or (Level At 1h > 6h)	Early Responders To Wounding	<ul style="list-style-type: none"> <li>Induction Of Key Signaling Pathways Within And Between Cells</li> <li>Modulation Of Wounding And Stress Induced Signal Transduction Pathways</li> <li>Specific Gene Transcription Initiation</li> <li>Induction Of Repair Processes Or Cell Death</li> <li>Reorientation Of Metabolism, Including Management Of Active Oxygen</li> <li>Movement Of Wound</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>Kinases And Phosphatases</li> <li>Jasmonic Acid, Salicylic Acid And Nitric Oxide Pathway Proteins.</li> <li>Glycohydrolases</li> <li>Dehydrins</li> <li>Rnases</li> <li>Metabolic Enzymes</li> <li>Nodulins</li> <li>Cell Division And Cell Wall Proteins</li> <li>Cold Response Proteins</li> <li>Lipoxygenase</li> <li>Jacalin</li> <li>Proteins To Detoxify Active Oxygen Species</li> <li>Systemin</li> </ul>

TRANSCRIPT LEVELS	TYPES OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		<p>Induced Signals Through Plant</p> <ul style="list-style-type: none"> <li>Synthesis Of Phytoalexins And Secondary Metabolites</li> </ul>	<ul style="list-style-type: none"> <li>Biosynthetic Enzymes</li> </ul>
Up Related Transcripts (Level At 1h < 6h)	<p>Delayed Responders</p> <p>Genes Involved In Wounding Response At Distant Sites From Wound.</p> <p>Genes Involved In Maintenance Of Wounding Response</p>	<ul style="list-style-type: none"> <li>Maintenance Of Defence Pathways</li> <li>Maintenance Of Reorientated Metabolism</li> <li>Maintenance Of Wound Response</li> <li>Programmed Cell Death In Selected Cells</li> <li>Reorientation Of Metabolism</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>Kinases And Phosphatases</li> <li>Jasmonic Acid, Salicylic Acid And Nitric Oxide Pathway Proteins</li> <li>Glycohydrolases</li> <li>Dehydrins</li> <li>Rnases</li> <li>Metabolic Enzymes</li> <li>Nodulins</li> <li>Cold Response Proteins</li> <li>Lipoxygenase</li> </ul>

TRANSCRIPT LEVELS	TYPES OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		<ul style="list-style-type: none"> <li>• Movement Of Wound Induced Signals Through Plant</li> <li>• Synthesis Of Phytoalexins And Secondary Metabolites</li> </ul>	<ul style="list-style-type: none"> <li>• Jacalin</li> <li>• Proteins To Detoxify Active Oxygen Species</li> <li>• Cell Division And Cell Wall Proteins</li> <li>• Systemin</li> <li>• Biosynthetic Enzymes</li> </ul>
Down – Regulated Transcripts (Level At 1h $\approx$ 6h) Or (Level At 6 Hr > 1h)	<ul style="list-style-type: none"> <li>• Early Responder Repressors Of Wounding Response State</li> <li>• Genes With Discontinued Expression Or UnsTable</li> </ul>	<ul style="list-style-type: none"> <li>• Negative Regulation Of Wounding Response Pathways Released</li> <li>• Changes In Pathways And Processes Operating In Cells</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Change In Protein Structure By Phosphory-Laton (Kinases) Or Dephos-Phorylation (Phosphatases)</li> <li>• Change In Chromatin Structure And Or Dna Topology</li> <li>• Local Changes In Regulatory Proteins, Metabolic Enzymes, Transporters Etc.</li> </ul>

TRANSCRIPT LEVELS	TYPES OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	mRNA Following Wounding		
Down – Regulated Transcripts (Level At 1hr > 6h)	Delayed Repressors Of Wounding Response State  Genes With Discontinued Expression Or Unstable mRNA Following Wounding	<ul style="list-style-type: none"> <li>Negative Regulation Of Wounding Response Pathways Released</li> <li>Change In Pathways And Process Operating In Cells</li> <li>Programmed Cell Death</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors, Phosphatases, Kinases</li> <li>Changes In Protein Complex Structures</li> <li>Chromatin Restructuring Proteins</li> <li>Local Changes In Regulatory Proteins, Metabolic Enzymes, Transporters Etc.</li> <li>Most Proteins In Selected Cells Undergoing Death</li> </ul>

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the wounding responsive genes when the desired sequence is operably linked to a promoter of a wounding responsive gene.

V.E. METHYL JASMONATE (JASMONATE) RESPONSIVE GENES, GENE  
COMPONENTS AND PRODUCTS

Jasmonic acid and its derivatives, collectively referred to as jasmonates, are naturally occurring derivatives of plant lipids. These substances are synthesized from linolenic acid in a lipoxygenase-dependent biosynthetic pathway. Jasmonates are signalling molecules which have been shown to be growth regulators as well as regulators of defense and stress responses. As such, jasmonates represent a separate class of plant hormones.

Changes in external or internal jasmonate concentration result in modulation of the activities of many genes and gene products. Examples of such "jasmonate responsive" genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield, especially when plants are growing in the presence of biotic or abiotic stresses. They were discovered and characterized from a much larger set of genes by experiments designed to find genes whose mRNA products changed in concentration in response to application of methyl jasmonate to plants.

Manipulation of one or more jasmonate responsive gene activities are useful to modulate the biological activities and/or phenotypes tested below. Jasmonate response genes and gene products can act alone or in combination. Useful combinations include jasmonate responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same co-regulated or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities. Such jasmonate responsive genes and gene products can function to either increase or dampen the phenotypes or activities below either in response to changes in jasmonate concentration or in the absence of jasmonate fluctuations. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108568, 108569, 108555). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.



MeJA genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### MeJA Genes Identified By Cluster Analyses Of Differential Expression

##### MeJA Genes Identified By Correlation To Genes That Are Differentially

##### Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of MeJA genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108568, 108569, 108555 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### MeJA Genes Identified By Correlation To Genes That Cause Physiological

##### Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of MeJA genes. A group in the MA\_clust is considered a MeJA pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### MeJA Genes Identified By Amino Acid Sequence Similarity

MeJA genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis MeJA genes. Groups of MeJA genes are identified in the Protein Group table. In this table, any protein group that comprises a

peptide ID that corresponds to a cDNA ID member of a MeJA pathway or network is a group of proteins that also exhibits MeJA functions/utilities.

Further, promoters of jasmonate responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by jasmonate or any of the following phenotypes or biological activities below.

5

V.E.1. USE OF JASMONATE RESPONSIVE GENES TO MODULATE  
PHENOTYPES:

Jasmonate responsive genes and their gene products can be used to alter or modulate one or more of the following phenotypes:

- Growth rate
- Whole Plant, including Height, Flowering Time, etc.
- Seedling
- Organ
- Coleoptile Elongation
- Young Leaves
- Roots
- Lateral Roots
- Tuber Formation
- Flowers
- Fruit
- Seeds
- Biomass
- Fresh and Dry Weight during any time in plant life, including maturation and senescence
- Number of Flowers
- Number of Seeds
- Secondary Metabolite Accumulation
- Alkaloids
- Anthocyanins
- Paclitaxel and Related Taxanes
- Rosmarinic
- Seed Yield
- Number, Size, Weight, Harvest Index
- Content and Composition, e.g., Amino Acid, Jasmonate, Oil, Protein, and
- Starch
- Fruit Yield

- Number, Size, Weight, Harvest Index, Post Harvest Quality
- Content and Composition e.g., Amino Acid, Carotenoid, Jasmonate, Protein, Starch
- Seed and Fruit Development
- Germination of Dormant and Non-Dormant Seeds
- Seed Viability
- Seed Reserve Mobilization
- Fruit Ripening
  - Initiation of the Reproductive cycle from a vegetative state
- Flower Development Time
- Insect Attraction for Fertilization
- Time to Fruit Maturity
- Senescence
- Fruits, Fruit Drop
- Leaves
- Stress and Disease Responses
- Drought
- Wounding
- UV damage
- Insect, Fungus, Virus, Worm damage

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the jasmonate responsive genes when the desired sequence is operably linked to a promoter of a jasmonate responsive gene.

To improve any of the phenotype(s) above, activities of one or more of the jasmonate responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed, for example, in accordance to citations described below.

#### V.E.2. USE OF JASMONATE-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES:

The activities of one or more of the jasmonate responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Turnover of proteins	- Induction of various proteases, ubiquitin and proteasome components and turnover of RNA polymerases and translation initiation factors - Reduction in many ribosomal proteins	This study. Standard biochemical assays.
Activation of nitrogen metabolism	Induction of glutamine synthetase, many aminotransferases, vegetative storage proteins	Crawford (1995) Plant Cell 7, 859-868 This study. Standard biochemical assays.
Lipid turnover	Induction of various lipases, desaturases, and reduction of lipid transfer protein mRNAs	This study. Standard biochemical assays.
Sugar metabolism	Induction of sugar transporters, UDP glucosyltransferases, other transferases	This study. Standard biochemical assays.

Glycolysis and central carbon metabolism	Induction of glycolytic related enzymes . Example, glucose 6-phosphate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglucumutase ATP synthase	This study. Standard biochemical assays.
Chlorosis	Degradation of Chlorophyll	Tsuchiya et al. (1999) Proc. Natl. Acad. Sci. U SA 96:15362-15367
	Inhibition of Photosynthesis Related Proteins	Reinbothe et al. (1993) J. Biol. Chem. 268, 10606-10611
Carbon Assimilation and turnover	Induction of chlorophyll ab binding protein precursor	Reinbothe et al. (1993) J. Biol. Chem. 268, 10606-10611
Jasmonate metabolism	Induction of lipid biosynthesis, myrosinase and jacalin	This study. Standard biochemical assays.
Jasmonate mediated signal transduction	Receptor binding	Cho and Pai (2000) Mol Cells 10, 317-324
	Protein kinases	Lee et al. (1998) Mol. Gen. Genet. 259, 516-522  Seo et al. (1999) Plant Cell 11, 289-298  Yoon et al. (1999) Plant Mol. Biol. 39, 991-1001

	Ubiquitination of Repressor Proteins	Xie et al. (1998) Science 280, 1091-1094
	Calcium Flux regulators	Bergey and Ryan (1999) Plant Mol. Biol. 40, 815-823
	Transcription Activators. Example- induction of various zinc finger, myb and AP-2 related factors	Xiang et al. (1996) Plant Mol. Biol. 32, 415-426  Menke et al. (1999) EMBO J. 18, 4455-4463
Response to Cell Membrane Damage	Lipid Peroxidation	Dubery et al. (2000) Mol. Cell Biol. Res. Commun. 3, 105-110
Cell Elongation	Inhibition of incorporation of Glucose into Cell Wall Saccharides	Burnett et al. (1993) Plant Physiol. 103, 41-48
Cell Organization and Division	- Reductions in tropomyosin related proteins and certain cyclins - Induction of actins and tubulins	Ishikawa et al. (1994) Plant Mol. Biol. 26, 403-414



Cell Wall Turnover and modulation	<ul style="list-style-type: none"> <li>- Induction of cell wall proteins, glycine-rich proteins, annexins, pectate lyase and pectin esterases</li> <li>- Reductions in various dehydrins and expansins</li> </ul>	<p>Creelman et al. (1992) Proc. Natl. Acad. Sci. USA 89, 4938-4941</p> <p>Garcia-Muniz et al. (1998) Plant Mol. Biol. 38, 623-632</p> <p>Norman et al (1999) Mol. Plant Microbe Interact. 12, 640-644</p>
Stress, Disease, and Pathogen Resistance	Induction of antifungal proteins, wounding responsive proteins, dehydrins, heat shock type proteins and elicitor response proteins	<p>Hildmann et al. (1992) Plant Cell 4, 1157-1170</p> <p>Reinbothe et al. (1994) Proc. Natl. Acad. Sci. USA 91, 7012-7016</p> <p>Moons et al. (1997) Plant Cell 9, 2243-2259</p> <p>Richard et al. (2000) Plant Mol. Biol. 43, 1-10</p> <p>Van Wees et al. (2000) Proc. Natl. Acad. Sci. USA 97, 8711-8716</p>

	Phytoalexin Biosynthesis	Creelman et al. (1992) Proc. Natl. Acad. Sci. USA 89, 4938-4941  Choi et al. (1994) Proc. Natl. Acad. Sci. USA 91, 2329-2333
	Biosynthesis of phenolics	Doares et al., (1995) Proc. Natl. Acad. Sci. USA 92, 4095-5098
	Production of Protease Inhibitors	Botella et al. (1996) Plant Physiol 112, 1201-1210
	Defense Gene Transcription in Response to UV	Mason et al. (1993) Plant Cell 5, 241-251  Schaller et al. (2000) Planta 210, 979-984
Secondary Metabolite biosynthesis	Fruit Carotenoid Composition	Czapski and Saniewski (1992) J. Plant Physiol. 139, 265-268
	Palitaxel and Related Taxanes	Yukimune et al. (1996) Nature Biotech. 14, 1129-1132

	Alkaloids	Aerts et al. (1994) Plant J. 4, 635-643 Geerlings et al. (2000) J. Biol. Chem. 275, 3051-3056
	Anthocyanins	Franceschi et al. (1991) Proc. Natl. Acad. Sci. USA 83, 6745-6749
	Rosmarinic	Mizukami et al., (1993) Plant Cell Reprod. 12, 706-709
	Activation of Ethylene- forming Enzyme and Production of Ethylene	Czapski and Saniewski (1992) J. Plant Physiol. 139, 265-268

Other biological activities that can be modulated by the jasmonate responsive genes and their products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Domain section of the Reference Tables.

Jasmonate responsive genes are characteristically differentially transcribed in response to fluctuating jasmonate levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various jasmonate responsive genes in the aerial parts of a seedling at 1 and 6 hours after being sprayed with Silwet L-77 solution enriched with methyl jasmonate as compared to seedlings sprayed with Silwet L-77 alone.

The data from this time course reveal a number of types of jasmonate responsive genes and gene products, including "early responders" and "delayed responders". Profiles of the individual kinds of jasmonate responsive genes are shown in the Table below, together with examples of the kinds of associated biological activities that are modulated when the activities of such genes vary.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
Upregulated genes (Level at 1 hour ≅ 6 hours). (Level at 1 hour > 6 hours)	Early Responders to Jasmonate	Binding and Perception of Jasmonate  Transduction of Jasmonate signal transduction response pathways  Initiation of Specific Gene Transcription to reorientate metabolism	Transcription Factors Transporters  Kinases, Phosphatases, Leucine-rich Repeat Proteins (LRRs), GTP- binding proteins (G- proteins), calcium- binding proteins and calcium responsive proteins  Proteases, lipases, glutamine synthetase (GS), arginase, aminotransferases, glycosyltransferases, sugar transporters, cell wall proteins, methyl transferases, glycolytic enzymes.
Upregulated genes (Level at 1 hour < 6 hours)	Delayed Jasmonate Responders	Maintenance of Metabolism under high Jasmonate	Enzymes of methyl jasmonate-induced pathways, including dehydrin, phytoalexin, phenolic, carotenoid, alkaloid and anthocyanin

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
		Jasmonate signal Transduction Response Pathways	biosynthesis.
		Gene Transcription to Reorientate Metabolism	Transcription factors, Transporters, Kinases and phosphatases
		Gene Transcription to Maintain Reorientated Metabolism	Proteases, Lipases, Glutaminae Synthetase, Arginase, Aminotransferases, Lipid Peroxidases, Glycosyltransferases, Sugar transporters, Cell Wall Proteins, Glycolytic Enzymes, Chlorophyll Binding Proteins
			Transcription factors, kinases, phosphatases, LRRs, G-proteins

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
		Reorient Cell Division and Cell Development	Actins, Tubulins, Myosins Cyclins, Cyclin-dependent Kinases (CDPKs)  Glycosyl Transferases, Glycosyl hydrolases, Expansins, Extensins, O- Methyl Transferases  Arabinogalactan- proteins (AGPs), Enzymes of Lipid Biosynthesis, Cutinase
Down regulated transcripts (level at 1 hour = 6 hours) ( level at 6 hours >1 hour)	Early responders of Jasmonate  Genes with discontinued expression or unsTable mRNA following Jasmonate uptake	Relese of Suppression of Jasmonate Induced Pathways  Reorientation of metabolism	Transcription Factors, Kinases, Phosphatases, LRRs, G-Proteins, Chromatin Restructuring proteins,  Ribosomal proteins, Translation Factors, Histones, RNA polymerases, Pectin esterase, Lipid transfer proteins

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	TYPE OF BIOLOGICAL ACTIVITY	EXAMPLES OF BIOCHEMICAL ACTIVITY
Down regulated transcripts (level at 1 hour > 6 hours)	Genes with Discontinued expression or UnsTable mRNA Following Jasmonate uptake	Negative Regulation of Jasmonate Induced Pathways Released.  Reorientation of metabolism	Transcription factors Kinases, Phosphatases Chromatin Restructuring Proteins, LRRs, G-proteins  Ribosomal proteins, Translation Factors, Histones RNA Polymerases, Cyclins Pectin esterase, Lipid Transfer Proteins

#### Use of Promoters of Jasmonate Responsive Genes

Promoters of Jasmonate responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Jasmonate responsive genes where the desired sequence is operably linked to a promoter of a Jasmonate responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.F. REACTIVE OXYGEN RESPONSIVE GENES, GENE COMPONENTS AND  
H<sub>2</sub>O<sub>2</sub> PRODUCTS

Often growth and yield are limited by the ability of a plant to tolerate stress conditions, including pathogen attack, wounding, extreme temperatures, and various other factors. To combat such conditions, plant cells deploy a battery of inducible defense responses, including triggering an oxidative burst. The burst of reactive oxygen intermediates occurs in time, place and strength to suggest it plays a key role in either pathogen elimination and/or subsequent signaling of downstream defense functions. For example, H<sub>2</sub>O<sub>2</sub> can play a key role in the pathogen resistance response, including initiating the hypersensitive response (HR). HR is correlated with the onset of systemic acquired resistance (SAR) to secondary infection in distal tissues and organs.

Changes in reactive oxygen, such as H<sub>2</sub>O<sub>2</sub> or O<sub>2</sub><sup>-</sup>, in the surrounding environment or in contact with a plant results in modulation of the activities of many genes and hence levels of gene products. Examples of such reactive oxygen responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. The genes were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to application of reactive oxygen, such as H<sub>2</sub>O<sub>2</sub>, to plants.

Manipulation of one or more reactive oxygen responsive gene activities are useful to modulate the following biological activities and/or phenotypes listed below. Reactive oxygen responsive genes and gene products can act alone or in combination. Useful combinations include reactive oxygen responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants.

Such reactive oxygen responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in reactive oxygen concentration or in the absence of reactive oxygen fluctuations. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the



experiment (see EXPT ID: 108582, 108583, 108537, 108538, 108558, and H2O2 (relating to SMD 7523)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Reactive Oxygen genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Reactive Oxygen Genes Identified By Cluster Analyses Of Differential Expression

##### Reactive Oxygen Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Reactive Oxygen genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108582, 108583, 108537, 108538, 108558, and H2O2 (relating to SMD 7523) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Reactive Oxygen Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Reactive Oxygen genes. A group in the MA\_clust is considered a Reactive Oxygen pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Reactive Oxygen Genes Identified By Amino Acid Sequence Similarity

Reactive Oxygen genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Reactive Oxygen genes. Groups of Reactive Oxygen genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Reactive Oxygen pathway or network is a group of proteins that also exhibits Reactive Oxygen functions/utilities.

Further, promoters of reactive oxygen responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by reactive oxygen or any of the following phenotypes or biological activities below.

#### V.F.1. USE OF REACTIVE OXYGEN RESPONSIVE GENES TO MODULATE PHENOTYPES

Reactive oxygen responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Pathogen Tolerance and/or Resistance
- Avr/R locus sensitive
- Non-Host sensitive
- HR
- SAR, e.g., where the reactive oxygen responsive gene and products are modulated in conjunction with any of the bacterial, fungal, virus, or other organism listed below
- Bacteria, resistance e.g. to *Erwinia stewartii*, *Pseudomonas syringae*, *Pseudomonas tabaci*, Stuart's wilt, etc.
- Fungal resistance including to downy mildews such as *Sclerophthora macrospora*, *Sclerophthora rayissiae*, *Sclerospora graminicola*, *Peronosclerospora sorghi*, *Peronosclerospora philippinensis*, *Peronosclerospora sacchari*, *Peronosclerospora maydis*; rusts such as *Puccinia sorphi*, *Puccinia polysora*, *Physopella zeae*, etc.; other fungal diseases such as *Cercospora zeae-maydis*, *Colletotrichum graminicola*, *Fusarium moniliforme*, *Exserohilum turcicum*, *Bipolaris maydis*, *Phytophthora parasitica*, *Peronospora tabacina*, *Septoria*, etc.;

- Virus or viroid resistance, e.g. to tobacco or cucumber mosaic virus, ringspot virus, necrosis virus, pelargonium leaf curl virus, red clover mottle virus, tomato bushy stunt virus, and like viruses;
- Insect resistance, such as to aphids e.g. *Myzus persicae*; beetles, beetle larvae; etc.
- nematodes, e.g. *Meloidogyne incognita*; lepidoptera, e.g. *Heliothus* spp. etc.
- Resistance Specifically in Primary or Secondary Leaves
- Stress Tolerance
- Winter Survival
- Cold Tolerance
- Heavy Metal Tolerance, such as Cadmium
- Physical Wounding;
- Increased Organelle Tolerance to Redox Stress, such as in Mitochondria, and chloroplasts
- Cell Death
- Apoptosis, including death of diseased tissue;
- Senescence;
- Fruit Drop;
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation;
- Number of Flowers, Seeds, Branches, and/or Leaves;
- Seed Yield, including Number, Size, Weight, and/or Harvest Index
- Fruit Yield, including Number, Size, Weight, and/or Harvest Index
- Plant Development
- Time to Fruit Maturity
- Cell Wall Strengthening and Reinforcement
- Plant Product Quality
  - Paper making quality
- Food additives
- Treatment of Indications modulated by Free Radicals
- Cancer

To regulate any of the phenotype(s) above, activities of one or more of the reactive oxygen

responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance to Alvarez et al., (1998) Cell 92: 773-784; Halhbrock and Scheel, (1989) Ann. Rev. Plant Physiol. Plant Mol. Biol. 40: 347-369; Lamb et al., (1997) Ann. Rev. Plant Mol. Biol. Plant Physio. 48: 251-275; Lapwood et al. (1984) Plant Pathol. 33: 13-20; Levine et al. (1996) Curr. Biol. 6: 427-437; McKersie et al., (2000) Plant Physiol. 122(4): 1427-1437; Olson and Varner (1993) Plant J. 4: 887-892; Pastore et al., (2000), FEBS Lett 470(1): 88-92; Pastori et al., (1997) Plant Physiol. 113: 411-418. Romero-Puertas et al., (1999) Free Radic. Res. 1999 31 Suppl: S25-31; Shirataki et al., Anticancer Res 20(1A): 423-426 (2000); Wu et al., (1995) Plant Cell 7: 1357-1368;

V.F.2. USE OF REACTIVE OXYGEN RESPONSIVE GENES TO  
MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the reactive oxygen responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations above and included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Reinforcement of Cell Walls	Modulation Of The Production Of Extracellular Proline-Rich Protein	Bradley et al. 1992. Cell 70, 21-30
	Modulation Of Lignification	Mansouri et al. (1999) Physiol Plant 106: 355-362
Stress, Disease, Pathogen Resistance and Wounding	Induction Of Pathogenesis Related Proteins, Phytoalexins And Many Defense Pathways.  Induction Of Detoxifying Enzymes Such As Glutathione S- Transferase And Ascorbate Peroxidase Disease Resistance	Chamnonngpol et.al.(1998) Proc. Nat.Acad Sci USA 12;95:5818-23. Davis et al. (1993) Phytochemistry 32: 607-611. Chen et.al. Plant J. (1996) 10:955-966 Gadea et.al.(1999) Mol Gen Genet 262:212-219  Wu et.al.(1995) Plant Cell 7: 1357-68

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Reactive Oxygen Generation Following Wounding And Changes In Physical Pressure	Orozco-Cardenas and Ryan (1999) Proc.Nat. Acad. Sci. USA 25;96:6553-7. Yahraus et al. (1995) Plant Physiol. 109: 1259-1266
	Modulation Of Genes Involved In Wound Repair And Cell Division	egendre et al. (1993) Plant hysiol. 102: 233-240
	Modulation Of Nitric Oxide Signaling	Delledonne et al. (1998) Nature 394: 585-588
	Salicyclic Acid Accumulation And Signaling	Durner and Klessig (1996) J.Biol.Chem. 271:28492-501
Programmed Cell Death	Induction Of Cell Death Pathway Genes	Levine et al. (1996) Curr. Biol. 6: 427-437. Reynolds et.al.(1998) Biochem.J. 330:115-20

Other biological activities that can be modulated by the reactive oxygen responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Reactive oxygen responsive genes are characteristically differentially transcribed in response to fluctuating reactive oxygen levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various reactive oxygen responsive genes in the aerial parts of a plant at 1 and 6 hours after the plant was sprayed with Silwett L-77 solution enriched with hydrogen peroxide as compared to plants sprayed with Silwett L-77 alone.

The data from this time course reveal a number of types of reactive oxygen responsive genes and gene products, including "early responders," and "delayed responders". Profiles of

individual reactive oxygen responsive genes are shown in the Table below together with examples of which associated biological activities are modulated when the activities of one or more such genes vary in plants.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSequence	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Upregulated transcripts (Higher at 1h Than 6h) (Level at 1 h $\cong$ 6h)	Early Responders To Reactive Oxygen	- Perceiving Reactive Oxygen  - Reactive Oxygen Response Transduction Pathways  - Initiating Specific Gene Transcription	-Transcription Factors -Kinases And Phosphatases -Transporters - Glutathione S-Transferase -Heat Shock Proteins -Salicylic Acid Response Pathway Proteins -Jasmonic Acid Pathway Proteins -Dehydrins -Peroxidases -Catalase -Proteases -Pathogen Response Proteins -Ca 2+ Channel Blockers -Phenylalanine Ammonia Lyase
Upregulated transcripts (Lower at 1h Than 6h)	Delayed Reactive Oxygen Responders	Maintenance Of Defence Pathways To Control Active Oxygen  Activation Of Cell	-Transcription Factors - Kinases And Phosphatases - Reactive Oxygen Scavenging Enzymes - Cell Wall And Cell Division/Growth Promoting Pathway Enzymes

484



	Are Minimised In Response To Reactive Oxygen	Activities Of Pathways Not Maintained Under Reactive Oxygen  Programmed Cell Death	- Many Proteins In Cells Undergoing Cell Death Or In Damaged Cells
--	--	--	--

Further, promoters of reactive oxygen responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by reactive oxygen or any of the following phenotypes or biological activities below.

V.G. SALICYLIC ACID RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Plant defense responses can be divided into two groups: constitutive and induced.

5 Salicylic acid (SA) is a signaling molecule necessary for activation of the plant induced defense system known as systemic acquired resistance or SAR. This response, which is triggered by prior exposure to avirulent pathogens, is long lasting and provides protection against a broad spectrum of pathogens. Another induced defense system is the hypersensitive response (HR). HR is far more rapid, occurs at the sites of pathogen (avirulent pathogens) entry and precedes SAR. SA is  
10 also the key signaling molecule for this defense pathway.

Changes in SA concentration in the surrounding environment or within a plant results in modulation of many genes and gene products. Examples of such SA responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. These  
15 genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to SA treatment.

While SA responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include  
20 different SA responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of SA responsive polynucleotides and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress and pathogen induced pathways,  
25 nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common and overlapping pathways.

Such SA responsive genes and gene products can function to either increase or dampen  
30 the above phenotypes or activities either in response to changes in SA concentration or in the absence of SA fluctuations. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or

AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108586, 108587, 108515, 108552, 108471, 108472, 108469, 108470, 107953, 107960, 108443, 108440, 108441, 108475, 108476). For transcripts that had higher levels in the samples than the control, a “+” is shown. A “-” is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

SA genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a “+” or “-” indication.

#### SA Genes Identified By Cluster Analyses Of Differential Expression

##### SA Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of SA genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108586, 108587, 108515, 108552, 108471, 108472, 108469, 108470, 107953, 107960, 108443, 108440, 108441, 108475, 108476 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### SA Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of SA genes. A group in the MA\_clust is considered a SA pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### SA Genes Identified By Amino Acid Sequence Similarity

SA genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis SA genes. Groups of SA genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID

that corresponds to a cDNA ID member of a SA pathway or network is a group of proteins that also exhibits SA functions/utilities.

Further, promoters of SA responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by SA or any of the following phenotypes or biological activities below.

#### V.G.1. USE OF SALICYLIC ACID-RESPONSIVE GENES TO MODULATE PHENOTYPES

SA responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Pathogen Tolerance and/or Resistance
- Avr/R locus Interactions
- Non-Host Interactions
- HR
- SAR, e.g., SA Responsive Genes And/Or Products In Conjunction With Any Of The Organisms Listed Below
- Resistance To Bacteria e.g. to *Erwinia stewartii*, *Pseudomonas syringae*, *Pseudomonas tabaci*, Stuart's wilt, etc.
- Resistance To Fungi e.g. to Downy Mildews Such As *Sclerophthora macrospora*, *Sclerophthora rayissiae*, *Sclerospora graminicola*, *Peronosclerospora sorghi*, *Peronosclerospora philippinensis*, *Peronosclerospora sacchari*, *Peronosclerospora maydis*; Rusts Such As *Puccinia sorphi*, *Puccinia polysora*, *Physopella zeae*, etc.; And To Other Fungal Diseases e.g. *Cercospora zeae-maydis*, *Colletotrichum graminicola*, *Fusarium moniliforme*, *Exserohilum turcicum*, *Bipolaris maydis*, *Phytophthora parasitica*, *Peronospora tabacina*, *Septoria*, etc.;
- Resistance To Viruses Or Viroids e.g., To Tobacco Or Cucumber Mosaic Virus, Ringspot Virus, Necrosis Virus, Pelargonium Leaf Curl Virus, Red Clover Mottle Virus, Tomato Bushy Stunt Virus, And Like Viruses;
- Resistance To Insects, Such As To Aphids e.g. *Myzus persicae*; to Beetles And Beetle Larvae; to lepidoptera larvae e.g. *Heliothus* etc.

- Resistance to Nematodes, e.g. *Meloidogyne incognita* etc
- Local Resistance In Primary (Infected) Or Secondary (Uninfected) Leaves
- Stress Tolerance
- Winter Survival
- Cold Tolerance
- Salt Tolerance
- Heavy Metal Tolerance, Such As Cadmium
- Tolerance To Physical Wounding;
- Increased Organelle Tolerance To Redox Stress, Such As In Mitochondria, And Chloroplasts
- Cell Death
- Programmed Cell Death, Including Death Of Diseased Tissue And During Senescence
- Fruit Drop
- Biomass
- Fresh And Dry Weight During Any Time In Plant Life, Such As Maturation
- Number Of Flowers, Seeds, Branches, And/Or Leaves
- Seed Yield, Including Number, Size, Weight, And/Or Harvest Index
- Fruit Yield, Including Number, Size, Weight, And/Or Harvest Index
- Plant Development
- Time To Fruit Maturity
- Cell Wall Strengthening And Reinforcement
- Plant Product Quality
  - Paper Making Quality
- Food Additives
- Treatment Of Indications Modulated By Free Radicals
- Cancer

To regulate any of the desired phenotype(s) above, activities of one or more of the SA responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a

plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Zhao et al. (1998, Plant Cell 10:359-70) and Alvarez et al. (1998, Cell 92: 733-84).

V.G.2. USE OF SALICYLIC ACID-RESPONSIVE GENES TO MODULATE  
BIOCHEMICAL ACTIVITIES

The activities of one or more of the SA responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATION INCLUDING ASSAYS
Protection From Microbial Pathogens	Systemic Acquired Resistance (SAR) - Phytoalexin Biosynthesis - PR Protein Biosynthesis Local Resistance Wound Response	Alvarez et al. (1998) Cell 92: 733-84 Lapwood et al. (1984) Plant Pathol. 33: 13-20 Davis et al. (1993) Phytochemistry 32: 607-11 Yahraus et al. (1995) Plant Physiol. 109: 1259-66
Cell Signaling	- Modulation Of Reactive Oxygen Signaling  - Modulation Of No Signaling	- Alvarez et al. (1998) Cell 92: 773-784 Delledonne et al. (1998) Nature 394: 585-588
Growth And Development	- Lignification	Redman et al. (1999) Plant Physiol. 119: 795-804

Other biological activities that can be modulated by the SA responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Salicylic acid responsive genes are characteristically differentially transcribed in response to fluctuating SA levels or concentrations, whether internal or external to an organism or cell.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various SA responsive genes in entire seedlings at 1 and 6 hours after the seedling was sprayed with a Hoagland's solution enriched with SA as compared to seedlings sprayed with Hoagland's solution only.

5

The data from this time course can be used to identify a number of types of SA responsive genes and gene products, including "early responders" and "delayed responders." Profiles of these different SA responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Upregulated Genes (Level At 1h $\approx$ 6h) Or (Level At 1h > 6h)	Early Responders To SA	- SA Perception - SA Uptake - Modulation Of SA Response Transduction Pathways	-Transcription Factors -Transporters, Kinases, Phosphatases, G- Proteins, LRR, DNA Remodelling Proteins
Upregulated Genes (Level At 1h < 6h)	Delayed Responders To SA	- Specific Defense gene Transcription Initiation (E.G. Pr Genes, Pal	-Proteases, PRProteins, Cellulases, Chitinases, Cutinases, Other Degrading Enzymes, Pal, Proteins Of Defense Pathways, Cell Wall Proteins  Epoxide Hydrolases, Methyl Transferases
Downregulated	- Early	- Negative Regulation	Transcription factors,



GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
(Level At 1h $\cong$ 6h) Or (Level At 6h > 1h)	Responder Repressors To SA - Genes With Discontinued Expression Or UnsTable mRNA In The Presence Of SA	Of SA Inducible Pathways Released	kinases, phosphatases, G- proteins, LRR, transporters, calcium binding proteins, chromatin remodelling protein
Down-Regulated Transcripts (Level At 1h > 6h)	- Delayed Responders To SA Metabolism - Genes With Discontinued Expression Or UnsTable mRNA In The Presence Of SA	Negative Regulation Of SA Inducible Pathways Released	Transcription Factors, Kinases, Phosphatases, G-Proteins, LRR, Transporters, Calcium Binding Proteins, Chromatin Remodelling Protein

Further, any desired sequence can be transcribed in similar temporal, tissue, or  
enviromentally specific patterns as the SA responsive genes when the desired sequence is operably  
linked to a promoter of a SA responsive gene.

V.H. NITRIC OXIDE RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

The rate-limiting element in plant growth and yield is often its ability to tolerate suboptimal or stress conditions, including pathogen attack conditions, wounding and the presence of various other factors. To combat such conditions, plant cells deploy a battery of inducible defense responses, including synergistic interactions between nitric oxide (NO), reactive oxygen intermediates (ROS), and salicylic acid (SA). NO has been shown to play a critical role in the activation of innate immune and inflammatory responses in animals. At least part of this mammalian signaling pathway is present in plants, where NO is known to potentiate the hypersensitive response (HR). In addition, NO is a stimulator molecule in plant photomorphogenesis.

Changes in nitric oxide concentration in the internal or surrounding environment, or in contact with a plant, results in modulation of many genes and gene products. Examples of such nitric oxide responsive genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. They were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to nitric oxide treatment.

While nitric oxide responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different nitric oxide responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of the levels of such proteins is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a nitric oxide responsive polynucleotide and/or gene product with other environmentally responsive polynucleotides is also useful because of the interactions that exist between hormone regulated pathways, stress pathways, pathogen stimulated pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways. The MA\_diff and/or AFLP\_diff MA\_diff

and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108584, 108585, 108526, 108527, 108559). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

NO genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### NO Genes Identified By Cluster Analyses Of Differential Expression

##### NO Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of NO genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108584, 108585, 108526, 108527, 108559 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### NO Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of NO genes. A group in the MA\_clust is considered a NO pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### NO Genes Identified By Amino Acid Sequence Similarity

NO genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis NO genes. Groups of NO

genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a NO pathway or network is a group of proteins that also exhibits NO functions/utilities.

Such nitric oxide responsive genes and gene products can function either to increase or dampen the above phenotypes or activities either in response to changes in nitric oxide concentration or in the absence of nitric oxide fluctuations. Further, promoters of nitric oxide responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by nitric oxide or any of the following phenotypes or biological activities below.

V.H.1. USE OF NITRIC OXIDE-RESPONSIVE GENES TO MODULATE  
PHENOTYPES :

Nitric oxide responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Stress Responses
- Mediation of response to stresses
- Disease resistance
- Growth
- Roots
- Stems
- Leaves
- Cells
- Promotes leaf cell elongation
- Biomass
- Fresh and Dry Weight during any time in plant life, such as at maturation;
- Size and/or Weight
- Flowers
- Seeds
- Branches

- Leaves
- Roots
- Development
- Seed Development
- Dormancy
- Control rate and timing of germination
- Prolongs seed storage and viability
- Senescence

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the nitric responsive genes when the desired sequence is operably linked to a promoter of a nitric responsive gene.

To regulate any of the desired phenotype(s) above, activities of one or more of the nitric oxide responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998) *Methods. Mol. Biol.* 82: 259-266 and/or screened for variants as described in Winkler et al. (1998) *Plant Physiol.* 118: 743-50 and visually inspected for the desired phenotype. Alternatively, plants can be metabolically and/or functionally assayed according to Beligni and Lamattina (2000) *Planta* 210: 215-21, Lapwood et al (1984) *Plant Pathol* 33: 13-20, and/or Brown and Botstein (1999) *Nature Genet.* 21: 33-37.

#### V.H.2. USE OF NITRIC OXIDE-RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES:

The activities of one or more of the nitric oxide responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Stress Response	-Programmed Cell Death  -Reactive Oxygen based Defence Pathways	Levine et al (1996) Curr. Biol 6: 427-37 Sellins and Cohen (1991) Radiat. Res. 126: 88-95 Kumar and Klessig (2000) Mol. Plant Microbe Interact. 13:347-351
Disease Resistance	-Microbial Pathogen resistance pathways        -Programmed Cell Death   -Cellular Protectant Gene expression  - Phytoalexin Biosynthesis	Lapwood et al (1984) Plant Pathol 33: 13-20 Kumar and Klessig (2000) Mol. Plant microbe interact.13: 347-351 Klessig et.al.(2000) Proc. Nat. Acad. Sci USA 97: 8849-8855 Delledonna et al(1998) Nature 394: 585-588  Levine et al (1996) Curr. Biol 6: 427-437 Sellins and Cohen (1991) Radiat. Res. 126: 88-95  Brown and Botstein (1999) Nat Genet 21: 33-37  Davis et al. (1993) Phytochemistry 32: 607-

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		611

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Signal Transduction	Regulation of hydrogen peroxide signaling	Wu et al. (1995) Plant Cell 7, 1357-1368
Reorientation of nitrogen metabolism	Induction of ribosomal proteins, asparagine synthesis, proteases, Rnases	This study. Standard assays for detection of changes
Reorientation of sugar and energy metabolism	Induction of sugar transporters, ATPases, glycohydrolases, and glycolytic enzymes, for example	This study. Standard assays for detection of changes

Other biological activities that can be modulated by the NO responsive genes and gene products are listed in the Reference Tables. Assays for detecting such biological activities are described in the Protein Domain table.

NO responsive genes are characteristically differentially transcribed in response to fluctuating NO levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various NO responsive genes in aerial tissues at 1 and 6 hours after a plant was sprayed with a Silwett L-77 solution enriched with 5 mM sodium nitroprusside, which is an NO donor. These changes are in comparison with plants sprayed with Silwett L-77 solution only.

The data from this time course can be used to identify a number of types of NO responsive genes and gene products, including “early responders” and “delayed responders” Profiles of these different nitric oxide responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.



GENE EXPRESSION LEVEL	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Upregulated genes (level at 1 hour $\approx$ 6 hours) (level at 1 hour > 6 hours)	Early responder repressors to NO	- NO Perception - NO Uptake - Modulation of NO Response Transduction Pathways  Specific Gene Transcription Initiation of Pathways to Optimize NO Response Pathways	-Transcription Factors  -Transporters -Pathogen responsive proteins, salicylic and jasmonate pathway proteins  -Proteins to provide defence against active oxygen e.g. glutathione transferase, ascorbate free radical reductase, ascorbate peroxidase, nitrilase, heat shock proteins -Proteins to reorient metabolism e.g. proteases, Rnases, proteasomes, asparagine synthetase, glycohydrolases, transporters -Proteins to inhibit transport of nitric oxide -Degradation enzymes

GENE EXPRESSION LEVEL	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Upregulated transcripts (level at 1 hour < 6 hours)	Delayed NO responders	<p>- Maintenance of metabolism in presence of High NO</p> <p>-Maintenance of disease defence pathways</p> <p>-Maintenance of pathways against reactive oxygen production</p> <p>Maintenance of different metabolic programs</p>	<p>- NO Metabolic Pathway enzymes</p> <p>-Pathogen responsive proteins, salicylic and jasmonate pathway proteins</p> <p>-Proteins to provide defence against active oxygen e.g. glutathione transferase, ascorbate free radical reductase, ascorbate peroxidase, nitrilase, heat shock proteins</p> <p>-Proteins to reorient and sustain metabolism e.g. proteases, RNases, pr oteasomes, asparagine synthetase, glycohydrol</p>
		Selective cell death	<p>ases, transporters,</p> <p>-Proteins to inhibit transport of NO</p> <p>-Degradation enzymes</p>
Down Regulated Transcripts (level at 1 hours $\approx$ 6)	-Early responders of NO utilization pathways	Negative regulation of NO utilization pathways released	<p>-Transcription factors</p> <p>-Kinases and phosphatases</p>



Use of Promoters of NO Responsive Genes

Promoters of NO responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the NO responsive genes where the desired sequence is operably linked to a promoter of a NO responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.I. OSMOTIC STRESS RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

The ability to endure and recover from osmotic and salt related stress is a major determinant of the geographical distribution and productivity of agricultural crops. Osmotic stress is a major component of stress imposed by saline soil and water deficit. Decreases in yield and crop failure frequently occur as a result of aberrant or transient environmental stress conditions even in areas considered suitable for the cultivation of a given species or cultivar. Only modest increases in the osmotic and salt tolerance of a crop species would have a dramatic impact on agricultural productivity. The development of genotypes with increased osmotic tolerance would provide a more reliable means to minimize crop losses and diminish the use of energy-costly practices to modify the soil environment.

Changes in the osmotic concentration of the surrounding environment or within a plant results in modulation of many genes and gene products. Examples of such osmotic stress responsive genes and gene products, including salt responsive genes, are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield.

While osmotic and/or salt stress responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different osmotic stress responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of an osmotic stress responsive polynucleotide and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in a common pathway.

Such osmotic and/or salt stress responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in osmotic concentration or in the absence of osmotic fluctuations. The MA\_diff and/or AFLP\_diff

MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: 108570, 108571, 108541, 108542, 108553, 108539, 108540). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Osmotic Stress genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Osmotic Stress Genes Identified By Cluster Analyses Of Differential Expression

##### Osmotic Stress Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Osmotic Stress genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID 108570, 108571, 108541, 108542, 108553, 108539, 108540 of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Osmotic Stress Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Osmotic Stress genes. A group in the MA\_clust is considered a Osmotic Stress pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Osmotic Stress Genes Identified By Amino Acid Sequence Similarity

Osmotic Stress genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Osmotic Stress genes. Groups of Osmotic Stress genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Osmotic Stress pathway or network is a group of proteins that also exhibits Osmotic Stress functions/utilities.

Further, promoters of osmotic stress responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by osmotic stress or any of the following phenotypes or biological activities below.

V.I.1. USE OF OSMOTIC STRESS RESPONSIVE GENES TO MODULATE  
PHENOTYPES

Osmotic stress responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
- Roots
- Stems
- Leaves
- Development
  - Cell Growth
    - DNA Synthesis and Cell Division
  - Seed Development
    - Desiccation tolerance
    - Dormancy
      - Control rate of Germination
      - Prolongs Seed Storage and Viability
  - Senescence
- Stress Responses
- Desiccation
- Drought

- Salt

To regulate any of the phenotype(s) above, activities of one or more of the osmotic stress responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to de Castro (1998, Phytochemistry 47: 689-694), Xu (1998, J Exp Bot 49: 573-582), Ausubel et al. (In: Current Protocols in Molecular Biology (1999) Volume 1, chapter 4, eds. Ausubel, Brent, Kingston, Moore, Seidman, Smith and Struhl, New York, NY) and De Castro et al. (2000, Plant Physiol 122: 327-36)

#### V.I.2. USE OF OSMOTIC STRESS RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the osmotic stress responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Growth And Differentiation	Regulation Of Osmolyte Synthesis	Yoshu et al. (1995) The Plant Journal 7: 751-60
	Regulation Of Glycolate Pathway And Photoinhibition Of Photosystem II In Response To Stress	Streb et al. (1993) Physiologia Plantarum. 88:590- 598



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Gene Regulation	Transcriptional Regulation Of Osmotic Stress Induced Proteins Through DNA Binding Proteins	Current Protocols in Molecular Biology / edited by Frederick M. Ausubel .. [et al.]. New York : Published by Greene Pub. Associates and Wiley- Interscience : J. Wiley, c1987
	Transcriptional Regulation Of Osmotic Stress Induced Proteins Through Protein Phosphorylation And Dephosphorylation	Jonak (1996) Proceedings of the National Academy of Sciences of the United States of America, 93: 11274-11279; Monroy, A.et al., (1998) Analytical Biochemistry 265: 183-185;
	Regulation Of Osmotic Stress Induced Gene Protein Accumulation By Protein Protein Intereaction Between Osmotic Stress Regulated Genes And Protein Phosphatase 2C	McCright (1998) IN: Methods in Molecular Biology; Protein phosphatase protocols; Ludlow (1998) Humana Press Inc. ; Suite 808, 999 Riverview Drive, Totowa, New Jersey 07512, USA. :263-277.
	Transcriptional Regulation Of Heat Induced Genes Through	Luo and Dean (1999) Journal of the

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Chromatin Remodeling	National Cancer Institute 91: 1288- 1294; Chromatin protocols (1999) edited by Peter B. Becker. Totowa, N.J. : Humana Press
	Activity Of Absciscic Acid Regulated DNA Binding Proteins	Gubler et al. (1999) Plant Journal 17: 1- 9
	Accumulation Of RNA Binding Proteins That Regulate Osmotic Stress	Sato (1995) Nucleic Acids Research 23: 2161-2167.
Stress Response	Synthesis And Metabolism Of Osmoprotectants Such As Betaine, Proline And Trehalase	Minocha et al. (1999) Plant Physiol and Biochem 37: 597- 603
	Regulation Of Sugar Transporters	Dejardin et al. (1999) Biochem J; 344 Pt 2:503-9
	Regulation Of Vacuolar Sodium/Proton Antiport Activity And The Detoxification Of Cations	Gaxiola et al. (1999) PNAS USA 96: 1480-1485
	Regulation Of Intracellular Na <sup>+</sup>	Espinoza-Ruiz et

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	And Li <sup>+</sup> Ion Concentrations	al. (1999) The Plant Journal 20: 529-539
	Regulation Of Universal Stress Protein Homologue Activity By Phosphorylation And Dephosphorylation.	Freestone et al. (1997) Journal of Molecular Biology, v. 274: 318-324
	Regulation/Maintenance Of Protein Stability During Thermal Stress	Walker (1996) Humana Press Inc. Suite 808, 999 Riverview Drive, Totowa, New Jersey 07512, USA
	Regulation Of Protein Degradation During Thermal Stress.	Vierstra (1996) Plant Molecular Biology, 32:275- 302.  Vierstra and Callis (1999) Plant Molecular Biology, 41:435-442
Signal Transduction	Activation Of Stress Response Genes	Xinong et al. (1999) The Plant Journal 19: 569-578
	Salt Tolerance	Piao (1999) Plant Physiol 19: 1527- 1534
	Calcium Mediated Stress Response	Subbaiah et al. (1994) Plant Physiology

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
		105:369-376 Kudla et al. (1999) PNAS USA 96: 4718-4723

Other biological activities that can be modulated by the osmotic stress responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Osmotic stress responsive genes are characteristically differentially transcribed in response to fluctuating osmotic stress levels or concentrations, whether internal or external to an organism or cell. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various osmotic stress responsive genes in aerial tissues of plants at 1 and 6 hours after the plants were sprayed with Hoagland's solution containing 20% PEG as compared to aerial tissues from plants sprayed with Hoagland's solution only.

The data from this time course can be used to identify a number of types of osmotic stress responsive genes and gene products, including "early responding," "sustained osmotic stress responders," "repressors of osmotic stress pathways" and "osmotic stress responders." Profiles of these different osmotic stress responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITIES OF GENE PRODUCTS
Up Regulated Transcripts (Level At 1 Hour $\equiv$ 6 Hours)	<ul style="list-style-type: none"> <li>• Early Responders To Osmotic Stress</li> </ul>	<ul style="list-style-type: none"> <li>• Osmotic Stress Perception</li> <li>• Osmolyte Uptake</li> <li>• Modulation Of</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Transcription Coactivators</li> </ul>



		Accumulation	<ul style="list-style-type: none"> <li>• Change In Chromatin Structure And/Or Localized Dna Topology</li> <li>• Modification Of Pre-Existing Translation Factors By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>• Synthesis Of New Translation Factors</li> <li>• Absciscic Acid Biosynthesis</li> </ul>
Up Regulated Transcripts (Level At 1 Hr < 6 Hr)	<ul style="list-style-type: none"> <li>• Sustained Osmotic Stress Responders</li> <li>• Repressor Of Osmotic Stress Pathways</li> <li>• Absciscic Acid Perception, Biosynthesis And Regulation</li> </ul>	<ul style="list-style-type: none"> <li>• Osmolyte Adjustment And Adaptation</li> <li>• Photosynthetic Activity Modification</li> <li>• Activation Of "Normal State" Biosynthesis Genes</li> <li>• Negative Regulation Of Osmotic Stress Pathways</li> <li>• Negative Regulation Of Absciscic Acid Biosynthesis</li> <li>• Activation Of Absciscic Acid Degradation</li> </ul>	<ul style="list-style-type: none"> <li>• Osmotic Stress Metabolic Pathways</li> <li>• Sugar Biosynthetic Pathways</li> <li>• Sugar Transporters</li> <li>• Transcription Factors</li> <li>• Transcription Coactivators</li> <li>• Membrane Transporters</li> <li>• Absciscic Acid Biosynthesis</li> </ul>

		Pathway	
		<ul style="list-style-type: none"> <li>Cell Wall Composition Modification</li> </ul>	
Down-Regulated Transcripts (Level At 1 Hr $\approx$ 6 Hr) (Level At 6 Hr $>$ 1 Hr)	Early Responder Repressors Of "Normal" State Of Metabolism  Negative Regulators Of Absciscic Acid Biosynthesis And Perception.  Positive Regulators Of "Normal State" Metabolic Pathways.	<ul style="list-style-type: none"> <li>Metabolic Repression</li> <li>Specific Gene Transcription Initiation</li> <li>Specific Gene Transcription Repression</li> <li>Translation Activation</li> <li>Translation Repression</li> <li>Absciscic Acid Degradation</li> <li>Protein Degradation</li> </ul>	<ul style="list-style-type: none"> <li>Transcription Factors</li> <li>Transcription Coactivators</li> <li>Protein Degradation</li> <li>Rna Binding Proteins</li> <li>Modification Of Protein Activity By Phosphatases, Kinases</li> <li>Activation Of Signaling Pathway By Calcium Binding Proteins,</li> <li>Modification Of Protein Activity By Protein-Protein Interaction</li> <li>Change In Chromatin Structure And/Or Localized Dna Topology</li> <li>Modification Of Pre-Existing</li> </ul>

			<p>Translation Factors By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</p> <ul style="list-style-type: none"> <li>• Synthesis Of New Translation Factors</li> </ul>
<p>Down-Regulated Transcripts (Level At 1 Hr &gt; 6 Hr)</p>	<p>Repressors Of "Normal" State Of Metabolism</p> <p>Genes With Discontinued Expression Or Unstable mRNA In Presence Of Osmotic Stress</p> <p>Repressor Of Osmotic Stress Pathways</p> <p>Repressors Of Abscisic Acid Biosynthesis, Perception And Regulation</p>	<ul style="list-style-type: none"> <li>• Osmotic Stress Adaptation</li> <li>• Negative Regulation Of Abscisic Acid Biosynthesis</li> <li>• Negative Regulation Of Osmotic Stress Response Pathways Genes</li> <li>• Osmolyte Synthesis And Osmolyte Cellular Partitioning Readjustment</li> <li>• Activation Of "Normal State" Metabolic Pathways</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription Factors</li> <li>• Transcription Coactivators</li> <li>• Protein Degradation</li> <li>• RNA Binding Proteins</li> <li>• Modification Of Protein Activity By Phosphatases, Kinases</li> <li>• Activation Of Signaling Pathway By Calcium Binding Proteins,</li> <li>• Modification Of Protein Activity By Protein-Protein Interaction</li> <li>• Change In Chromatin Structure And/Or Localized DNA</li> </ul>



			<p>Topology</p> <ul style="list-style-type: none"> <li>• Modification Of Pre-Existing Translation Factors By Phosphorylation (Kinases) Or Dephosphorylation (Phosphatases)</li> <li>• Synthesis Of New <ul style="list-style-type: none"> <li>◦ Translation Factors</li> </ul> </li> <li>• Sugar Biosynthetic Pathways</li> <li>• Sugar Transporters</li> </ul>
--	--	--	---

Further, any desired sequence can be transcribed in similar temporal, tissue, or enviromentally specific patterns as the osmotic stress responsive genes when the desired sequence is operably linked to a promoter of an osmotic stress responsive gene.

V.J. ALUMINUM RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Aluminum is toxic to plants in soluble form ( $Al^{3+}$ ). Plants grown under aluminum stress have inhibited root growth and function due to reduced cell elongation, inhibited cell division and metabolic interference. As an example, protein inactivation frequently results from displacement of the  $Mg^{2+}$  cofactor with aluminum. These types of consequences result in poor nutrient and water uptake. In addition, because stress perception and response occur in the root apex, aluminum exposure leads to the release of organic acids, such as citrate, from the root as the plant attempts to prevent aluminum uptake.

The ability to endure soluble aluminum is a major determinant of the geographical distribution and productivity of agricultural crops. Decreases in yield and crop failure frequently occur as a result of aberrant, hot conditions even in areas considered suitable for the cultivation of a given species or cultivar. Only modest increases in the aluminum tolerance of crop species would have a dramatic impact on agricultural productivity. The development of genotypes with increased aluminum tolerance would provide a more reliable means to minimize crop losses and diminish the use of costly practices to modify the environment.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by simultaneously hybridizing two differentially labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing 10,000 non-redundant ESTs, selected from 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases over the controls were identified, compared to the Ceres full-length cDNA and genomic sequence databanks, and identical Ceres clones identified. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which are aluminum response responsive genes.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Aluminum (relating to SMD 7304,

SMD 7305)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Aluminum genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Aluminum Genes Identified By Cluster Analyses Of Differential Expression

##### Aluminum Genes Identified By Correlation To Genes That Are Differentially

##### Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Aluminum genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Aluminum (relating to SMD 7304, SMD 7305) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Aluminum Genes Identified By Correlation To Genes That Cause Physiological

##### Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Aluminum genes. A group in the MA\_clust is considered a Aluminum pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Aluminum Genes Identified By Amino Acid Sequence Similarity

Aluminum genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Aluminum genes. Groups of Aluminum genes are identified in the Protein Group table. In this table, any protein

group that comprises a peptide ID that corresponds to a cDNA ID member of a Aluminum pathway or network is a group of proteins that also exhibits Aluminum functions/utilities.

V.J.1. USE OF ALUMINUM RESPONSE GENES TO MODULATE  
PHENOTYPES

Changes in aluminum concentrations in a plant's surrounding environment results in modulation of many genes and gene products. Examples of such aluminum response genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield.

While aluminum responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different aluminum responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of a aluminum responsive polynucleotide and/or gene product with another environmentally responsive polynucleotide is also useful because of the interactions that exist between hormone regulated pathways, stress pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in a common pathway.

Such aluminum responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either

- in response to changes in aluminum concentration or
- in the absence of aluminum fluctuations.

More specifically, aluminum responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
- Roots

Inhibition of root elongation

- Stems

- Leaves
- Whole Plant
- Development
  - Cell Growth
    - Elongation
    - Division
  - Mediates response to oxidative stress, calcium-mediated defense, antioxidant defense and pathogenesis

To produce the desired phenotype(s) above, one or more of the aluminum response genes or gene products can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, *Methods. Mol. Biol.* 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Li and Fleming (1999, *FEBS Lett* 461: 1-5), Delhaize et al. (1999, *J Biol Chem* 274: 7082-8), Sigimoto and Sakamoto (1997, *Genes Genet Syst* 72: 311-6), Esaki et al. (2000, *Plant Physiol* 122: 657-65), Leonard and Gerber (1988, *Mutat Res* 196: 247-57), Baisakhi et al. (2000, *Mutat Res* 465: 1-9), Ma (2000, *Plant Cell Physiol* 41: 383-90) and Koyama et al. (1999, *Plant Cell* 40: 482-8)

Alternatively, the activities of one or more of the aluminum responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the

Table below:

GENERAL CATEGORY	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	ASSAY
Cell Growth and Development	-Phospholipase D (PLD) activity	Toda et al. (1999) Biosci Biotechnol

	<p>-Regulation of Phosphatidylserine Synthase (PSS)</p> <p>-Cell wall strengthening</p>	<p>Biochem 63: 210-212</p> <p>Hamel et al. (1998) Planta 205: 531-38</p>
Stress Response	<p>-Regulation of oxidative stress</p> <p>-Regulation of antioxidant defense and DNA repair</p> <p>-Secretion of Organic Acids (e.g. maleate, citrate) from root apex</p> <p>-Ca<sup>2+</sup>-mediated Defense Responses Against Low pH</p>	<p>Esaki et al. (2000) Plant Physiol 122: 657-655</p> <p>Baisakhi et al. (2000) Mutat Res 465: 1-9</p> <p>Koyama et al. (1999) Plant Cell 40: 482-8</p> <p>Plieth et al. (1999) Plant J 18: 634-50</p>
Signaling	<p>-H<sup>+</sup> transport</p> <p>-Auxin transport</p>	<p>Degenhardt et al. (1988) Plant Physiol 117: 19-27</p> <p>Rashotte et al. (2000) Plant Physiol 122: 481-90</p>

Other biological activities that can be modulated by aluminum response genes and their products are listed in the REFERENCE Table. Assays for detecting such biological activities are described in the Protein Domain table.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	responders to aluminum application	<ul style="list-style-type: none"> <li>Aluminum perception</li> <li>Aluminum uptake and transport</li> <li>Aluminum metabolism</li> <li>Synthesis of secondary metabolites and/or proteins</li> <li>Modulation of aluminum response transduction pathways</li> <li>Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>Transporters</li> <li>Metabolic enzymes</li> <li>Change in cell membrane structure and potential</li> <li>Kinases and phosphatases</li> <li>Transcription activators</li> <li>Change in chromatin structure and/or localized DNA topology</li> </ul>
Down-regulated transcripts	<p>responder to aluminum repressors of aluminum state of metabolism</p> <p>Genes with</p>	<ul style="list-style-type: none"> <li>Negative regulation of aluminum pathways</li> <li>Changes in pathways and</li> </ul>	<ul style="list-style-type: none"> <li>Transcription factors</li> <li>Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> </ul>



TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	discontinued expression or unstable mRNA in presence of aluminum	<ul style="list-style-type: none"> <li>processes operating in cells</li> <li>Changes in other metabolisms than aluminum</li> </ul>	<ul style="list-style-type: none"> <li>Change in chromatin structure and/or DNA topology</li> <li>Stability of factors for protein synthesis and degradation</li> <li>Metabolic enzymes</li> </ul>

#### Use of Promoters of Aluminum Responsive Genes

Promoters of Aluminum responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Aluminum responsive genes where the desired sequence is operably linked to a promoter of a Aluminum responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.K. CADMIUM RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Cadmium (Cd) has both toxic and non-toxic effects on plants. Plants exposed to non-toxic concentrations of cadmium are blocked for viral disease due to the inhibition of systemic movement of the virus. Surprisingly, higher, toxic levels of Cd do not inhibit viral systemic movement, suggesting that cellular factors that interfere with the viral movement are triggered by non-toxic Cd concentrations but repressed in high Cd concentrations. Furthermore, exposure to non-toxic Cd levels appears to reverse posttranslational gene silencing, an inherent plant defense mechanism. Consequently, exploring the effects of Cd exposure has potential for advances in plant disease control in addition to soil bio-remediation and the improvement of plant performance in agriculture.

Changes in cadmium concentrations in a plant's surrounding environment results in modulation of many genes and gene products. Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by simultaneously hybridizing two differentially labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing some 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in plants treated with 10  $\mu$ M cadmium compared with untreated plants were identified, compared to the Ceres full length cDNA and genomic sequence databanks, and the equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which represent cadmium responsive genes.

Examples of such cadmium responsive genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield.

While cadmium responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations

include different cadmium responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants. In addition, the combination of a cadmium responsive polynucleotide and/or gene product with other environmentally responsive polynucleotides is also useful because of the interactions that exist between, for example, stress and pathogen induced pathways, nutritional pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Cadium (relating to SMD 7427, SMD 7428)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Cadium genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Cadium Genes Identified By Cluster Analyses Of Differential Expression

#### Cadium Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Cadium genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Cadium (relating to SMD 7427, SMD 7428) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

5                   Cadium Genes Identified By Correlation To Genes That Cause Physiological  
Consequences

10                   Additionally, the differential expression data and the phenotypic observations can  
15                   be merged to identify pathways or networks of Cadium genes. A group in the MA\_clust is  
                    considered a Cadium pathway or network if the group comprises a cDNA ID that also appears in  
                    Knock-in or Knock-out tables that causes one or more of the phenotypes described in section  
                    above.

20                   Cadium Genes Identified By Amino Acid Sequence Similarity

25                   Cadium genes from other plant species typically encode polypeptides that share  
                    amino acid similarity to the sequences encoded by corn and Arabidopsis Cadium genes. Groups  
                    of Cadium genes are identified in the Protein Group table. In this table, any protein group that  
                    comprises a peptide ID that corresponds to a cDNA ID member of a Cadium pathway or  
                    network is a group of proteins that also exhibits Cadium functions/utilities.

30                   Such cadmium responsive genes and gene products can function to either increase or  
                    dampen phenotypes or activities either in response to changes in cadmium concentration or in the  
                    absence of cadmium fluctuations. Further, promoters of cadmium responsive genes, as described  
                    in the Reference tables, for example, are useful to modulate transcription that is induced by  
                    cadmium or any of the following phenotypes or biological activities below.

35                   V.K.1. USE OF CADMIUM RESPONSIVE GENES, GENE COMPONENTS  
                    AND PRODUCTS TO MODULATE PHENOTYPES

                    Cadmium responsive genes and gene products are useful to or modulate one or more of  
                    the following phenotypes:

- 40                   • Growth  
                    • Roots

- Initiation and maintenance of cell division
- Stems
- Leaves
- Development
- Mitochondria
- Post-embryonic root meristem development
- Senescence
- Stress Response
- Modulation of Jasmonic Acid and other stress control pathways
- Metabolic detoxification
- Heavy metals
  - Plant and Seed Yield; Fruit Yield

Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the cadmium responsive genes when the desired sequence is operably linked to a promoter of a cadmium responsive gene.

To regulate any of the phenotype(s) above, activities of one or more of the cadmium responsive genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998) Methods. Mol. Biol. 82:259-266) and/or screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Ghoshroy et al. (1998, Plant J 13: 591-602), Citovsky et al. (1998, Plant J 16: 13-20), Clemens et al. (1999, EMBO J 18: 3325-33), Chen et al. (2000, Chemosphere 41: 229-34), Xian and Oliver (1998, Plant Cell 10: 1539-90), Romero-Peurtas et al. (1999, Free Rad Res 31: S25-31), Gaur and Noraho (1995, Biomed Environ Sci 8: 202-10), Thomine et al. (2000, PNAS USA 97: 4991-6), Howden et al. (1995, Plant Physiol 107: 1067-73), Kessler and Brand (1994, Eur J Biochem 225: 907-22) and Vernoux et al. (2000, Plant Cell 12: 97-110).

V.K.2. USE OF CADMIUM-RESPONSIVE GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the cadmium responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below.

5 Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation and Development	Root Growth Initiation and maintenance of cell division Resistance to Cadmium- inhibition of root growth	Thomine et al. (2000) PNAS USA <u>97</u> : 4991-6 Vernoux et al. (2000) Plant Cell <u>12</u> : 97-110
Metabolism	Cadmium sensing	Howden et al. (1995) Plant Physiol <u>107</u> : 1067-73
	Cadmium uptake and transport Decreased cadmium transport Phytoremediation	Gaur and Noraho (1995) Biomed Environ Sci <u>8</u> : 202-10 Thomine et al. (2000) PNAS USA <u>97</u> : 4991-6
	Inhibition of oxidative phosphorylation	Kessler and Brand (1994) Eur. Biochem <u>225</u> : 907-22
Plant Defenses	Viral resistance Inhibition of systemic movement of virus Block of viral disease	Ghoshroy et al. (1998) Plant J <u>13</u> : 591-602
	Detoxification of heavy metals	Clemens et al. (1999) EMBO J <u>18</u> : 3325-33
	Enhanced stress resistance	Romero-Peurtas et al. (1999) Free Rad Res <u>31</u> : S25-31
	Cadmium resistance	Xiang and Oliver (1998) Plant

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	via modulation of jasmonic acid signaling pathway	Cell 10: 1539-90
Signaling	Relief of post-translational gene silencing	Citovsky et al. (1998) Plant J 16: 13-20

Other biological activities that can be modulated by the cadmium responsive genes and gene products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Cadmium responsive genes are characteristically differentially transcribed in response to fluctuating cadmium levels or concentrations, whether internal or external to an organism or cell. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various cadmium responsive genes following treatment with 10  $\mu$ M cadmium, relative to untreated plants.

Profiles of some cadmium responsive genes are shown in the Table below together with examples of the kinds of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	Responders to cadmium Application  Genes induced by cadmium	Cadmium perception Cadmium uptake and transport Cadmium metabolism Synthesis of secondary metabolites and/or proteins Modulation of cadmium response	Transporters Metabolic enzymes Change in cell membrane structure and potential Kinases and Phosphatases Transcription activators Change in chromatin structure and/or localized



TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
		transduction pathways Specific gene transcription initiation Genes involved in inhibiting systemic movement of plant viral RNA Genes involved in post translational gene silencing	DNA topology RNA binding proteins
Down-regulated transcripts	Responders to cadmium  Genes repressed by cadmium  Genes with discontinued expression or unsTable mRNA in presence of cadmium	Negative regulation of cadmium pathways released  Changes in pathways and processes operating in cells Changes in metabolism other than cadmium pathways Genes involved in facilitating systemic movement of plant viral RNA Genes involved in promoting post translational gene silencing	Transcription factors Change in protein structure by phosphorylation (kinases) or Dephosphoryaltion (phosphatases) Change in chromatin structure and/or DNA topology Factors for protein synthesis and degradation Metabolic enzymes RNA binding proteins

Use of Promoters of Cadmium Responsive Genes

Promoters of Cadmium responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Cadmium responsive genes where the desired sequence is operably linked to a promoter of a Cadmium responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.L. DISEASE RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Often growth and yield are limited by the ability of a plant to tolerate stress conditions, including pathogen attack. To combat such conditions, plant cells deploy a battery of inducible defense responses, including the triggering of an oxidative burst and the transcription of pathogenesis-related protein (PR protein) genes. These responses depend on the recognition of a microbial avirulence gene product (avr) by a plant resistance gene product (R), and a series of downstream signaling events leading to transcription-independent and transcription-dependent disease resistance responses. Reactive oxygen species (ROS) such as  $H_2O_2$  and NO from the oxidative burst plays a signaling role, including initiation of the hypersensitive response (HR) and induction of systemic acquired resistance (SAR) to secondary infection by unrelated pathogens. PR proteins are able to degrade the cell walls of invading microorganisms, and phytoalexins are directly microbicidal.

The presence of an avirulent pathogen and/or changes in the concentrations of  $O_2^-$ ,  $H_2O_2$  and NO in the environment surrounding a plant cell modulate the activities of many genes and, therefore, the levels of many gene products. Examples of tobacco mosaic virus (TMV) responsive genes and gene products, many of them operating through an ROS signaling system, are shown in The Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield. The genes were discovered and characterized from a much larger set by experiments designed to find genes whose mRNA products changed in response to application of TMV to plants.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing some 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in response to TMV infection over the non infected controls were identified, compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff table(s) report(s) the

results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which represent disease responsive genes.

Manipulation of one or more disease responsive gene activities are useful to modulate the biological processes and/or phenotypes listed below. Disease responsive genes and gene products can act alone or in combination. Useful combinations include disease responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants.

Such disease responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in active oxygen concentration or in the absence of active oxygen fluctuations. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Disease (relating to SMD 7342, SMD 7343)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Disease genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

#### Disease Genes Identified By Cluster Analyses Of Differential Expression

##### Disease Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Disease genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Disease (relating to SMD 7342, SMD 7343) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

#### Disease Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Disease genes. A group in the MA\_clust is considered a Disease pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

#### Disease Genes Identified By Amino Acid Sequence Similarity

Disease genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Disease genes. Groups of Disease genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Disease pathway or network is a group of proteins that also exhibits Disease functions/utilities.

Further, promoters of disease responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by disease or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the disease responsive genes when the desired sequence is operably linked to a promoter of a disease responsive gene.

#### V.L.1. USE OF DISEASE RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE PHENOTYPES

Disease responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Pathogen Tolerance and/or Resistance
- Avr/R locus interactions
- Non-Host interactions

- HR
- SAR
- Resistance to bacteria e.g. to *Erwinia stewartii*, *Pseudomonas syringae*, *Pseudomonas tabaci*, Stuart's wilt, etc.
- Resistance to fungi e.g. to downy mildews such as *Sclerophthora macrospora*, *Sclerophthora rayissiae*, *Sclerospora graminicola*, *Peronosclerospora sorghi*, *Peronosclerospora philippinensis*, *Peronosclerospora sacchari*, *Peronosclerospora maydis*; rusts such as *Puccinia sorphi*, *Puccinia polysora*, *Physopella zae*, etc.; and to other fungal diseases e.g. *Cercospora zae-maydis*, *Colletotrichum graminicola*, *Fusarium moniliforme*, *Exserohilum turcicum*, *Bipolaris maydis*, *Phytophthora parasitica*, *Peronospora tabacina*, *Septoria*, etc.;
- Resistance to viruses or viroids e.g. to tobacco or cucumber mosaic virus, ringspot virus, necrosis virus, pelargonium leaf curl virus, red clover mottle virus, tomato bushy stunt virus, and like viruses;
- Resistance to insects, such as to aphids e.g. *Myzus persicae*; to beetles and beetle larvae; to lepidoptera larvae e.g. *Heliothus* etc.
- Resistance to Nematodes, e.g. *Meloidogyne incognita* etc
- Local resistance in primary (infected) or secondary (uninfected) leaves
- Stress Tolerance
- Winter Survival
- Cold Tolerance
- Salt tolerance
- Heavy Metal Tolerance, such as Cadmium
- Tolerance to Physical Wounding;
- Increased Organelle Tolerance to Redox Stress, such as in Mitochondria, and chloroplasts
- Cell Death
- Programmed cell death, including death of diseased tissue and during senescence
- Fruit Drop
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation

- Number of Flowers, Seeds, Branches, and/or Leaves
- Seed Yield, including Number, Size, Weight, and/or Harvest Index
- Fruit Yield, including Number, Size, Weight, and/or Harvest Index
- Plant Development
- Time to Fruit Maturity
- Cell Wall Strengthening and Reinforcement
- Plant Product Quality
- Paper making quality
- Food additives
- Treatment of Indications modulated by Free Radicals
- Cancer
- Kinds of low molecular weight compounds such as phytoalexins
- Abundance of low molecular weight compounds such as phytoalexins
- Other phenotypes based on gene silencing

To regulate any of the phenotype(s) above, activities of one or more of the disease responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance to Alvarez et al., (1998) Cell 92: 773-784; Halbrock and Scheel, (1989) Ann. Rev. Plant Physiol. Plant Mol. Biol. 40: 347-369; Lamb et al., (1997) Ann. Rev. Plant Mol. Biol. Plant Physiol. 48: 251-275; Lapwood et al. (1984) Plant Pathol. 33: 13-20; Levine et al. (1996) Curr. Biol. 6: 427-437; McKersie et al., (2000) Plant Physiol. 122: 1427-1437; Olson and Varner (1993) Plant J. 4: 887-892; Pastore et al., (2000), FEBS Lett 470: 88-92; Pastori et al., (1997) Plant Physiol. 113: 411-418; Romero-Puertas et al., (1999) Free Radic. Res. 1999 31 Suppl: S25-31; Shirataki et al., Anticancer Res 20: 423-426 (2000); Wu et al., (1995) Plant Cell 7: 1357-1368.

V.L.2. USE OF DISEASE RESPONSIVE GENES, GENE COMPONENTS  
AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the disease responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations above and included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Resistance to Pathogens	Induction of ROS signaling pathways	Wu et.al.(1995) Plant Cell 7: 1357-68
	Modulation of nitric oxide signaling	Delledonne et al. (1998) Nature 394: 585-588
	Induction of PR proteins, phytoalexins, and defense pathways	Chamnongpol et.al.(1998) Proc. Nat.Acad Sci USA 12;95:5818-23. Davis et al. (1993) Phytochemistry 32: 607-611
	Induction of cellular protectant genes such as glutathione S-transferase (GST) and ascorbate peroxidase	Chen et.al. Plant J. (1996) 10:955-966 Gadea et.al.(1999) Mol Gen Genet 262:212-219 Wu et.al.(1995) Plant Cell 7: 1357-68
	ROS levels following wounding and changes in physical pressure	Orozco-Cardenas and Ryan (1999) Proc.Nat. Acad. Sci. USA 25;96:6553-7. Yahraus et al. (1995) Plant Physiol. 109: 1259-1266



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	Salicylic acid levels and signaling	Durner and Klessig (1996) J.Biol.Chem. 271:28492-501
Responses to Wounding	Expression of genes Involved in wound repair and cell division	Legendre et al. (1993) Plant Physiol. 102: 233-240
Responses to Environmental Stress	Expression of genes involved in responses to drought, cold, salt, heavy metals	Shi et al. (2000) Proc. Natl. Acad. Sci. USA 97:6896-6901
Reinforcement of Cell Walls	Modulation of the Production of Extracellular Proline-Rich Protein	Bradley et al. (1992) Cell 70, 21- 30
	Modulation of Lignification	Mansouri et al. (1999) Physiol. Plant 106: 355-362
Programmed Cell Death	Induction of PCD activating genes	Levine et al. (1996) Curr. Biol. 6: 427-437. Reynolds et.al. (1998) Biochem.J. 330:115-20
	Suppression of PCD suppressing genes	Pennell and Lamb (1997) Plant Cell 9, 1157-1168

Other biological activities that can be modulated by the disease responsive genes and their products are listed in the Reference Table. Assays for detecting such biological activities are described in the Protein Domain table.

5

Disease responsive genes are characteristically differentially transcribed in response to fluctuating levels of disease. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the changes in transcript levels of various disease responsive genes in the aerial parts of a plant 3 days after the plant was sprayed with a suspension of TMV relative to control plants sprayed with water.

The data from this experiment reveal a number of types of disease responsive genes and gene products, including "early responders," and "delayed responders". Profiles of individual disease responsive genes are shown in the Table below with examples of which associated biological activities are modulated when the activities of one or more such genes vary in plants.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSequence	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Upregulated transcripts	Early Responders to Pathogens	ROS Perception and Response	Transcription factors, kinases, phosphatases, GTP- binding proteins (G- proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodeling proteins
		Initiation of Gene Transcription	Glutathione S-transferase (GST), heat shock proteins, salicylic acid (SA) response pathway proteins, jasmonate response pathway proteins, dehydrins, peroxidases, catalases
	Delayed Responders to Pathogens	Initiation of Defence Gene Transcription	Proteases, pathogen response (PR) proteins, cellulases, chitinases, cutinases, glucanases, other degrading enzymes, calcium channel blockers, phenylalanine ammonia lyase, proteins of defense pathways, cell wall proteins including proline rich

			proteins and glycine rich proteins, epoxide hydrolase, methyl transferases
		Activation of cell death pathways	Transcription factors kinases, phosphatases, DNA surveillance proteins, p53, proteases, endonucleases, GTP-binding proteins (G-proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, mitochondrial and chloroplast energy related proteins, ribosome inactivating proteins
		Initiation of Cellular Protectant Gene Transcription	Reactive oxygen scavenging enzymes, GST, catalase, peroxidase, ascorbate oxidase
Downregulated transcripts	Early responders to pathogens	Negative regulation of pathogen inducible pathways released	Transcription factors, kinases, phosphatases, GTP-binding proteins (G-proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins
	Genes repressed by TMV	Negative regulation of ROS inducible	Transcription factors, kinases, phosphatases, GTP-

		pathways released	binding proteins (G-proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins
	Delayed Responders to Pathogens	Negative regulation of pathogen inducible pathways released	Transcription factors, kinases, phosphatases, GTP-binding proteins (G-proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins
	Genes repressed by TMV	Negative regulation of genes suppressing programmed cell death released	Transcription factors, kinases, phosphatases, GTP-binding proteins (G-proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins

#### Use of Promoters of Disease Responsive Genes

Promoters of Disease responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Disease responsive genes where the desired sequence is operably linked to a promoter of a Disease responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such

promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins,  
or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

003555, 002401

V.M. DEFENSE (LOL2) RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS

Often growth and yield are limited by the ability of a plant to tolerate stress conditions, including pathogen attack. To combat such conditions, plant cells deploy a battery of inducible defense responses, including the triggering of an oxidative burst and the transcription of pathogenesis-related protein (PR protein) genes. Reactive oxygen species (ROS) such as H<sub>2</sub>O<sub>2</sub> and NO from the oxidative burst play a signaling role, including initiation of the hypersensitive response (HR) and induction of systemic acquired resistance (SAR) to secondary infection by unrelated pathogens. Some PR proteins are able to degrade the cell walls of invading microorganisms, and phytoalexins are directly microbicidal. Other defense related pathways are regulated by salicylic acid (SA) or methyl jasmonate (MeJ).

These responses depend on the recognition of a microbial avirulence gene product (avr) by a plant resistance gene product (R), and a series of downstream signaling events leading to transcription-independent and transcription-dependent disease resistance responses. Current models suggest that R- gene-encoded receptors specifically interact with pathogen-encoded ligands to trigger a signal transduction cascade. Several components include ndr1 and eds1 loci. NDR1, EDS1, PR1, as well as PDF1.2, a MeJ regulated gene and Nim1, a SA regulated gene, are differentially regulated in plants with mutations in the LOL2 gene.

LOL2 shares a novel zinc finger motif with LSD1, a negative regulator of cell death and defense response. Due to an alternative splice site the LOL2 gene encodes two different proteins, one of which contains an additional, putative DNA binding motif. Northern analysis demonstrated that LOL2 transcripts containing the additional DNA binding motif are predominantly upregulated after treatment with both virulent and avirulent *Pseudomonas syringae* pv *maculicola* strains. Modulation in this gene can also confer enhanced resistance to virulent and avirulent *Peronospora parasitica* isolates.

Examples of LOL2 responsive genes and gene products are shown in the Reference, Sequence, Protein Group, Protein Group Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff and MA\_clust tables. These genes and/or products are responsible for effects on traits such as plant vigor, disease resistance, and seed yield. The genes were discovered and characterized from a much larger set by microarray experiments designed to find genes whose mRNA products changed when the LOL2 gene was mutated in plants.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing some about 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in plants with the LOL2 mutation versus wildtype were obtained. Specifically, the plant line lol-2-2 tested, a loss of function mutation. The ESTs were compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which represent LOL2 responsive genes.

Manipulation of one or more LOL2 responsive gene activities are useful to modulate the biological processes and/or phenotypes listed below. LOL2 responsive genes and gene products can act alone or in combination. Useful combinations include LOL2 responsive genes and/or gene products with similar transcription profiles, similar biological activities, or members of the same or functionally related biochemical pathways. Whole pathways or segments of pathways are controlled by transcription factor proteins and proteins controlling the activity of signal transduction pathways. Therefore, manipulation of such protein levels is especially useful for altering phenotypes and biochemical activities of plants.

Such LOL2 responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in active LOL2 gene or in the absence. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: lol2 (relating to SMD 8031, SMD 8032)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.



Defense genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

5        Defense Genes Identified By Cluster Analyses Of Differential Expression

Defense Genes Identified By Correlation To Genes That Are Differentially Expressed

10        As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

15        A pathway or network of Defense genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID lol2 (relating to SMD 8031, SMD 8032) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Defense Genes Identified By Correlation To Genes That Cause Physiological Consequences

20        Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Defense genes. A group in the MA\_clust is considered a Defense pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

25        Defense Genes Identified By Amino Acid Sequence Similarity

30        Defense genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Defense genes. Groups of Defense genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Defense pathway or network is a group of proteins that also exhibits Defense functions/utilities.

Further, promoters of LOL2 responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by LOL2 responsive genes or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or enviromentally specific patterns as the LOL2 responsive genes when the desired sequence is operably linked to a promoter of a LOL2 responsive gene.

V.M.1. USE OF LOL2 RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE PHENOTYPES

Lol2 Responsive Genes And Gene Products Are Useful To Or Modulate One Or More Of

The Following Phenotypes:

- Pathogen Tolerance and/or Resistance
- Avr/R locus interactions
- Non-Host interactions
- HR
- SAR, e.g., disease responsive genes acting in conjunction with infection with any of the organisms listed below
- Resistance to bacteria e.g. to *Erwinia stewartii*, *Pseudomonas syringae*, *Pseudomonas tabaci*, Stuart's wilt, etc.
- Resistance to fungi e.g. to downy mildews such as *Sclerophthora macrospora*, *Sclerophthora rayissiae*, *Sclerospora graminicola*, *Peronosclerospora sorghi*, *Peronosclerospora philippinensis*, *Peronosclerospora sacchari*, *Peronosclerospora maydis*; rusts such as *Puccinia sorphi*, *Puccinia polysora*, *Physopella zeae*, etc.; and to other fungal diseases e.g. *Cercospora zeae-maydis*, *Colletotrichum graminicola*, *Fusarium moniliforme*, *Exserohilum turcicum*, *Bipolaris maydis*, *Phytophthora parasitica*, *Peronospora tabacina*, *Septoria*, etc.;
- Resistance to viruses or viroids e.g. to tobacco or cucumber mosaic virus, ringspot virus, necrosis virus, pelargonium leaf curl virus, red clover mottle virus, tomato bushy stunt virus, and like viruses;
- Resistance to insects, such as to aphids e.g. *Myzus persicae*; to beetles and beetle larvae; to lepidoptera larvae e.g. *Heliothis* etc.
- Resistance to Nematodes, e.g. *Meloidogyne incognita* etc
- Local resistance in primary (infected) or secondary (uninfected) leaves
- Stress Tolerance
- Winter Survival
- Cold Tolerance
- Salt tolerance

- Heavy Metal Tolerance, such as Cadmium
- Tolerance to Physical Wounding;
- Increased Organelle Tolerance to Redox Stress, such as in Mitochondria, and chloroplasts
- Cell Death
- Programmed cell death, including death of diseased tissue and during senescence
- Fruit Drop
- Biomass
- Fresh and Dry Weight during any time in plant life, such as maturation
- Number of Flowers, Seeds, Branches, and/or Leaves
- Seed Yield, including Number, Size, Weight, and/or Harvest Index
- Fruit Yield, including Number, Size, Weight, and/or Harvest Index
- Plant Development
- Time to Fruit Maturity
- Cell Wall Strengthening and Reinforcement
- Plant Product Quality
- Paper making quality
- Food additives
- Treatment of Indications modulated by Free Radicals
- Cancer
- Kinds of low molecular weight compounds such as phytoalexins
- Abundance of low molecular weight compounds such as phytoalexins
- Other phenotypes based on gene silencing

To regulate any of the phenotype(s) above, activities of one or more of the LOL2 responsive genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be screened for variants as in Winkler et al. (1998) Plant Physiol 118: 743-50 and assayed, for example, in accordance to Alvarez et al., (1998) Cell 92: 773-784; Halbrock and Scheel, (1989) Ann. Rev. Plant Physiol. Plant Mol. Biol. 40: 347-369; Lamb et al., (1997) Ann. Rev. Plant Mol. Biol. Plant Physiol. 48:

251-275; Lapwood et al. (1984) Plant Pathol. 33: 13-20; Levine et al. (1996) Curr. Biol. 6: 427-437; McKersie et al., (2000) Plant Physiol. 122: 1427-1437; Olson and Varner (1993) Plant J. 4: 887-892; Pastore et al., (2000), FEBS Lett 470: 88-92; Pastori et al., (1997) Plant Physiol. 113: 411-418; Romero-Puertas et al., (1999) Free Radic. Res. 1999 31 Suppl: S25-31; Shirataki et al.,  
5 Anticancer Res 20: 423-426 (2000); Wu et al., (1995) Plant Cell 7: 1357-1368.

### V.M.2. USE OF DEFENSE RESPONSIVE GENES TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the defense (LOL2) responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities are documented and can be measured according to the citations above and included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Resistance To Pathogens	Induction Of ROS Signaling Pathways	Wu et.al.(1995) Plant Cell 7: 1357-68
	Modulation Of Nitric Oxide Signaling	Delledonne et al. (1998) Nature 394: 585-588
	Induction Of PR Proteins, Phytoalexins, And Defense Pathways	Chamnongpol et.al.(1998) Proc. Nat.Acad Sci USA 12;95:5818-23. Davis et al. (1993) Phytochemistry 32: 607-611
	Induction Of Cellular Protectant Genes Such As Glutathione S-Transferase (GST) And Ascorbate Peroxidase	Chen et.al. Plant J. (1996) 10:955-966 Gadea et.al.(1999) Mol Gen Genet 262:212-219 Wu et.al.(1995) Plant Cell 7: 1357-68

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	ROS Levels Following Wounding And Changes In Physical Pressure	Orozco-Cardenas and Ryan (1999) Proc.Nat. Acad. Sci. USA 25;96:6553-7. Yahraus et al. (1995) Plant Physiol. 109: 1259-1266
	Salicyclic Acid Levels And Signaling	Durner and Klessig (1996) J.Biol.Chem. 271:28492-501
Responses To Wounding	Expression Of Genes Involved In Wound Repair And Cell Division	Legendre et al. (1993) Plant Physiol. 102: 233-240
Responses To Environmental Stress	Expression Of Genes Involved In Responses To Drought, Cold, Salt, Heavy Metals	Shi et al. (2000) Proc. Natl. Acad. Sci. USA 97:6896-6901
Reinforcement Of Cell Walls	Modulation Of The Production Of ExtracTable Proline-Rich Protein	Bradley et al. (1992) Cell 70, 21-30
	Modulation Of Lignification	Mansouri et al. (1999) Physiol. Plant 106: 355-362
Programmed Cell Death	Induction Of Pcd Activating Genes	Levine et al. (1996) Curr. Biol. 6: 427-437. Reynolds et.al. (1998) Biochem.J. 330:115-20
	Suppression Of PCD Suppressing Genes	Pennell and Lamb (1997) Plant Cell 9, 1157-1168

Other biological activities that can be modulated by the LOL2 responsive genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

LOL2 responsive genes are characteristically differentially transcribed in response to fluctuating levels of disease. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various LOL2 responsive genes in the lol-2 line versus control plants.

5 The data from this experiment reveal a number of types of LOL2 responsive genes and gene products. Profiles of individual LOL2 responsive genes are shown in the Table below with examples of which associated biological activities are modulated when the activities of one or more such genes vary in plants.

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSequence	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
Upregulated transcripts	Early Responders to the LOL2 Mutation	ROS Perception and Response	Transcription factors, kinases, phosphatases, GTP- binding proteins (G- proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodeling proteins
		Initiation of Gene Transcription	Glutathione S-transferase (GST), heat shock proteins, salicylic acid (SA) response pathway proteins, jasmonate response pathway proteins, dehydrins, peroxidases, catalases
	Delayed Responders to the LOL2	Initiation of Defence Gene Transcription	Proteases, pathogen response (PR) proteins, cellulases, chitinases,

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSequence	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
	Mutation		cutinases, glucanases, other degrading enzymes, calcium channel blockers, phenylalanine ammonia lyase, proteins of defense pathways, cell wall proteins including proline rich proteins and glycine rich proteins, epoxide hydrolase, methyl transferases
		Activation of cell death pathways	Transcription factors kinases, phosphatases, DNA surveillance proteins, p53, proteases, endonucleases, GTP-binding proteins (G- proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, mitochondrial and chloroplast energy related proteins, ribosome inactivating proteins
		Initiation of Cellular Protectant Gene Transcription	Reactive oxygen scavenging enzymes, GST, catalase, peroxidase, ascorbate

003662 002404



GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSEQUENCE	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
			oxidase
Downregulated transcripts	Early Responders to the LOL2 Mutation	Negative regulation of LOL2 Mutation inducible pathways released	Transcription factors, kinases, phosphatases, GTP- binding proteins (G- proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins
	Genes Repressed by the LOL2 Mutation	Negative regulation of ROS inducible pathways released	Transcription factors, kinases, phosphatases, GTP- binding proteins (G- proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins
	Delayed Responders to the LOL2 Mutation	Negative regulation of LOL2 Mutation inducible pathways released	Transcription factors, kinases, phosphatases, GTP- binding proteins (G- proteins), leucine rich repeat proteins (LRRs), transporters, calcium binding proteins, chromatin remodelling proteins
	Genes	Negative Regulation Of	Transcription Factors,

GENE EXPRESSION LEVELS	FUNCTIONAL CATEGORY OF GENE	PHYSIOLOGICAL CONSequence	EXAMPLES OF BIOCHEMICAL ACTIVITY OF GENE PRODUCTS
	Repressed By The LOL2 Mutation	Genes Suppressing Programmed Cell Death Released	Kinases, Phosphatases, GTP-Binding Proteins (G- Proteins), Leucine Rich Repeat Proteins (Lrrs), Transporters, Calcium Binding Proteins, Chromatin Remodelling Proteins

#### Use of Promoters of Defense Responsive Genes

Promoters of Defense responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Defense responsive genes where the desired sequence is operably linked to a promoter of a Defense responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.N. IRON RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Iron (Fe) deficiency in humans is the most prevalent nutritional problem worldwide today. Increasing iron availability via diet is a sustainable malnutrition solution for many of the world's nations. One-third of the world's soils, however, are iron deficient. Consequently, to form a food-based solution to iron malnutrition, we need a better understanding of iron uptake, storage and utilization by plants. Furthermore, exposure to non-toxic Fe levels appears to affect inherent plant defense mechanisms. Consequently, exploring the effects of Fe exposure has potential for advances in plant disease resistance in addition to human nutrition.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by simultaneously hybridizing two differentially labeled fluorescent FeNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing 10,000 non-redundant ESTs, selected from 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases over the controls were identified, compared to the Ceres full length FeNA and genomic sequence databanks, and identical Ceres clones identified. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones that are up or down regulated over controls, thereby indicating the Ceres clones which are iron responsive genes.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Iron (relating to SMD 7114, SMD 7115, SMD 7125)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Iron genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

Iron Genes Identified By Cluster Analyses Of Differential Expression

Iron Genes Identified By Correlation To Genes That Are Differentially

Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Iron genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Iron (relating to SMD 7114, SMD 7115, SMD 7125) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Iron Genes Identified By Correlation To Genes That Cause Physiological

Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Iron genes. A group in the MA\_clust is considered a Iron pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Iron Genes Identified By Amino Acid Sequence Similarity

Iron genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Iron genes. Groups of Iron genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Iron pathway or network is a group of proteins that also exhibits Iron functions/utilities.

V.N.1. USE OF IRON RESPONSIVE GENES TO MODULATE PHENOTYPES

Iron responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth

- Roots
- Root hair formation
- Stems
- Leaves
- Development
- Senescence
- Plant nutrition
- Uptake and assimilation of organic compounds
- Uptake and assimilation of inorganic compounds
- Animal (including human) nutrition
- Improved dietary mineral nutrition
- Stress Response
- Metabolic detoxification
- Heavy metals

To improve any of the phenotype(s) above, activities of one or more of the iron responsive genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, Methods. Mol. Biol. 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Schmidt et al. (2000, Plant Physiol 122:1109-18), Meagher (2000) Current Opinion in Plant Biology 3: 153-62), Deak (1999, Nature Biotechnology (1999, Nature Biotechnology 17: 192-96), Wei and Theil (2000, J. Biol Chem 275: 17488-93) and Vansuyt et al. (1997, FEBS Letters 410: 195-200).

#### V.N.2. USE OF IRON-RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the iron responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below.

Such biological activities can be measured according to the citations included in the  
Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation and Development	<ul style="list-style-type: none"> <li>Root Growth <ul style="list-style-type: none"> <li>Initiation of root hairs</li> </ul> </li> </ul>	Robinson et al. (1999) Nature 397: 694-97
Metabolisms	<ul style="list-style-type: none"> <li>Iron sensing</li> <li>Iron uptake and transport <ul style="list-style-type: none"> <li>-decreased iron transport</li> <li>-phytoremediation</li> </ul> </li> </ul>	Thomine et al. (2000) PNAS USA 97: 4991-6 Thomine et al. (2000) PNAS USA 97: 4991-6 Zhu (1999) Plant Physiol 119: 73-79
Plant Defenses	<ul style="list-style-type: none"> <li>Protection from oxidative damage</li> </ul>	Deak (1999) Nature Biotechnology 17: 192-6
Signaling	<ul style="list-style-type: none"> <li>Specific gene transcription gene silencing</li> </ul>	Brand and Perrimon (1993) Development 118: 401-415

Other biological activities that can be modulated by the iron responsive genes and gene products are listed in the REFERENCE Table. Assays for detecting such biological activities are described in the Protein Domain table.

Iron responsive genes are characteristically differentially transcribed in response to fluctuating iron levels or concentrations, whether internal or external to an organism or cell. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various iron responsive genes.

The microarray comparison consists of probes prepared from root RNA of A. thaliana (Columbia) seedlings grown under iron-sufficient conditions and seedlings grown under iron-deficient. The data from this experiment reveal a number of types genes and gene products.

Profiles of these different iron responsive genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	responders to iron application	<ul style="list-style-type: none"> <li>• Iron perception</li> <li>• Iron uptake and transport</li> <li>• Iron metabolism</li> <li>• Synthesis of secondary metabolites and/or proteins</li> <li>• Modulation of iron response transduction pathways</li> <li>• Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic enzymes</li> <li>• Change in cell membrane structure and potential</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> </ul>
Down-regulated transcripts	<p>responder to iron repressors of iron state of metabolism</p> <p>Genes with discontinued expression or</p>	<ul style="list-style-type: none"> <li>• Negative regulation of iron pathways</li> <li>• Changes in pathways and processes</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> </ul>



TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	unstable mRNA in presence of iron	operating in cells <ul style="list-style-type: none"> <li>• Changes in other metabolisms than iron</li> </ul>	<ul style="list-style-type: none"> <li>• Change in chromatin structure and/or DNA topology</li> <li>• Stability of factors for protein synthesis and degradation</li> <li>• Metabolic enzymes</li> </ul>

#### Use of Promoters of Iron Responsive Genes

Promoters of Iron responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Iron responsive genes where the desired sequence is operably linked to a promoter of a Iron responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.O. SHADE RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Plants sense the ratio of Red (R) : Far Red (FR) light in their environment and respond differently to particular ratios. A low R:FR ratio, for example, enhances cell elongation and favors flowering over leaf production. The changes in R:FR ratios mimic and cause the shading response effects in plants. The response of a plant to shade in the canopy structures of agricultural crop fields influences crop yields significantly. Therefore manipulation of genes regulating the shade avoidance responses can improve crop yields. While phytochromes mediate the shade avoidance response, the down-stream factors participating in this pathway are largely unknown. One potential downstream participant, ATHB-2, is a member of the HD-Zip class of transcription factors and shows a strong and rapid response to changes in the R:FR ratio. ATHB-2 overexpressors have a thinner root mass, smaller and fewer leaves and longer hypocotyls and petioles. This elongation arises from longer epidermal and cortical cells, and a decrease in secondary vascular tissues, paralleling the changes observed in wild-type seedlings grown under conditions simulating canopy shade. On the other hand, plants with reduced ATHB-2 expression have a thick root mass and many larger leaves and shorter hypocotyls and petioles. Here, the changes in the hypocotyl result from shorter epidermal and cortical cells and increased proliferation of vascular tissue. Interestingly, application of auxin is able to reverse the root phenotypic consequences of high ATHB-2 levels, restoring the wild-type phenotype. Consequently, given that ATHB-2 is tightly regulated by phytochrome, these data suggest that ATHB-2 may link the auxin and phytochrome pathways in the shade avoidance response pathway.

Changes in R:FR ratios promote changes in gene expression. Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by hybridizing labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The US Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing about 10,000 non-redundant ESTs, selected from about 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases in plants given 4 hours of FR rich light after growth in high R:FR light compared with the controls of

plants grown in high R:FR light only, were identified, compared to the Ceres full length cDNA and genomic sequence databanks, and equivalent Ceres clones identified. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s) report(s) the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby  
5 indicating the Ceres clones which are shade avoidance responsive genes.

Examples of far red light induced, shade avoidance responsive genes and gene products are shown in the Reference and Sequence Tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed yield.

While far red light, shade avoidance responsive polynucleotides and gene products can act alone, combinations of these polynucleotides also affect growth and development. Useful combinations include different shade avoidance responsive polynucleotides and/or gene products that have similar transcription profiles or similar biological activities, and members of the same or similar biochemical pathways. In addition, the combination of a shade avoidance responsive polynucleotide and/or gene product with another environmentally responsive polynucleotides is also useful because of the interactions that exist between hormone regulated pathways, stress and pathogen induced pathways, nutritional pathways, light induced pathways and development. Here, in addition to polynucleotides having similar transcription profiles and/or biological activities, useful combinations include polynucleotides that may have different transcription profiles but which participate in common or overlapping pathways.  
10  
15

Such far red light induced shade avoidance responsive genes and gene products can function to either increase or dampen the above phenotypes or activities either in response to changes in far red light or in the absence of far red light fluctuations. The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Shade (relating to SMD 8130, SMD 7230)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.  
20  
25

Shade genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication  
30

Shade Genes Identified By Cluster Analyses Of Differential Expression

Shade Genes Identified By Correlation To Genes That Are Differentially

Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Shade genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Shade (relating to SMD 8130, SMD 7230) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Shade Genes Identified By Correlation To Genes That Cause Physiological

Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Shade genes. A group in the MA\_clust is considered a Shade pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Shade Genes Identified By Amino Acid Sequence Similarity

Shade genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Shade genes. Groups of Shade genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Shade pathway or network is a group of proteins that also exhibits Shade functions/utilities.

Further, promoters of shade avoidance responsive genes, as described in the Reference tables, for example, are useful to modulate transcription that is induced by shade avoidance or any of the following phenotypes or biological activities below. Further, any desired sequence can be transcribed in similar temporal, tissue, or environmentally specific patterns as the shade

avoidance responsive genes when the desired sequence is operably linked to a promoter of a circadian (clock) responsive gene.

V.O.1. USE OF FAR RED RESPONSIVE, SHADE AVOIDANCE RESPONSE  
GENES TO MODULATE PHENOTYPES

High FR:R, shade avoidance responsive genes and gene products can be used to alter or modulate one or more of the following phenotypes:

- Growth
- Roots
- Elongation
- Lateral root formation
- Stems
- Elongation
- Expansion
- Leaves
- Expansion
- Carotenoid composition
- Development
- Cell
- Growth
- Elongation
- Photosynthetic apparatus
- Efficiency
- Flower
- Flowering time
- Fruit
- Seed
- Dormancy
- Control rate and timing of germination
- Prolongs seed storage and viability
- Inhibition of hydrolytic enzyme synthesis
- Seed and Fruit yield
- Senescence
- Abscission

- Leaf fall
- Flower longevity
- Differentiation
- Vascularization
- Shade (Avoidance) Responses in plant architecture

To regulate any of the phenotype(s) above, activities of one or more of the High FR: R light, shade avoidance responsive genes or gene products can be modulated and the plants tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, *Methods. Mol. Biol.* 82:259-266) and/or screened for variants as in Winkler et al. (1998) *Plant Physiol* 118: 743-50 and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Carabelli et al. (1996, *PNAS USA* 93: 3530-3535), Aguirrezabal and Tardieu (1996, *J Exp Bot* 47: 411-20), Heyer et al. (1995, *Plant Physiol* 109: 53-61), Garcia-Plazaola et al. (1997, *J Exp Bot* 48: 1667-74), Schwanz et al. (1996, *J Exp Bot* 47L 1941-50), Koehne et al. (1999, *Biochem Biophys Acta* 1412:94-107), Melis (1984, *J Cell Biochem* 24: 271-85), Steindeler et al. (1999, *Development* 126: 4235-45), Cruz (1997, *J Exp Bot* 48: 15-24), Stephanou and Manetas (1997, *J Exp Bot* 48: 1977-85), Grammatikopoulos et al. (1999, *J Exp Bot* 50:517-21), Krause et al. (1999, *Plant Physiol* 121: 1349-58), Aukerman et al. (1997, *Plant Cell* 9: 1317-26), Wagner et al. (1997, *Plant Cell* 9: 731-43), Weinig (2000) *Evolution Int J Org Evolution* 54: 124-26), Cocburn et al. (1996, *J Exp Bot* 47: 647-53), Devlin et al. (1999, *Plant Physiol* 119: 909-15), Devlin et al. (1998, *Plant Cell* 10: 1479-87), Finlayson et al. (1998, *Plant Physiol* 116: 17-25), Morelli and Ruberti (2000, *Plant Physiol* 122: 621-26), Aphalo et al. (1999, *J Exp Bot* 50: 1629-34), Sims et al. (1999, *J Exp Bot* 50: 50: 645-53) and Ballare (1999, *Trends Plant Sci* 4: 97-102).

V.O.2. USE OF FAR RED LIGHT, SHADE AVOIDANCE RESPONSIVE  
GENES TO MODULATE BIOCHEMICAL ACTIVITIES

5 The activities of one or more of the far red light, shade avoidance responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:



PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Cell Growth and Differentiation	Cell Elongation	Carabelli et al. (1996) PNAS USA 93: 3530-35
	Leaf Expansion	Heyer et al. (1995) Plant Physiol 109: 53-61
Photosynthesis	Development of Photosynthetic Apparatus	Jagtap et al. (1998) J Exp Bot 49: 1715-21 Melis (1984) J Cell Biochem 24: 271-285 McCain (1995) Biophys J 69: 1105- 10
	Carotenoid Composition	Garcia-Plazaola et al (1997) J Exp Bot 48: 1667-74
Carbon/Nitrogen Metabolism	Carbon and Nitrogen Assimilation	Cruz (1997) J Exp Bot 48: 15-24
Far red light, shade avoidance response binding by transcription factors		Newton AL, Sharpe BK, Kwan A, Mackay JP, Crossley M. J Biol Chem. 2000May19;275(20):15128- 34; Lopez Ribera I, Ruiz-Avila L, Puigdomenech P. Biochem Biophys Res Commun. 1997 Jul 18;236(2):510-6; de Pater S, Greco V, Pham K, Memelink J, Kijne J. Nucleic Acids Res. 1996 Dec 1; 24(23):4624-31.

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Signaling	UV Light Perception	Stephanou and Manetas (1997) J Exp Bot 48: 1977-85
	Far-red/Red Light Perception	Aukerman et al. (1997) Plant Cell 9: 1317-26 Wagner et al. (1997) Plant Cell 9: 731-43
	Interaction of "Shade Factor" with Ethylene Production/Transduction	Finlayson et al. (1998) Plant Physiol 116: 17-25
	Interaction of "Shade Factor" with Auxin Production/Transduction	Reed et al. (1998) Plant Physiol 118: 1369-78
	Plant to Plant signalling	Sims et al. (1999) J Exp Bot 50: 645-53

Other biological activities that can be modulated by shade avoidance response genes and their products are listed in the REF TABLES. Assays for detecting such biological activities are described in the Protein Domain table.

- 5 High FR:R, shade avoidance responsive genes are differentially transcribed in response to high FR:R ratios. The microarray comparison to reveal such genes consisted of probes prepared from RNA isolated from the aerial tissues of *A. thaliana* (Columbia) two-week old seedlings grown in high R:FR ratios compared to seedlings grown in high R:FR ratios followed by 4 hours of FR-rich light treatment. The data from this experiment reveal a number of types genes and
- 10 gene products and examples of the classes of genes are given in the Table below.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	<p>Responders to high FR:R light ratios</p> <p>Genes induced by high FR:R light ratio</p>	<ul style="list-style-type: none"> <li>Far red light perception</li> <li>Metabolism affected by far red light</li> <li>Synthesis of secondary metabolites and/or proteins</li> <li>Modulation of high FR:R light ratio transduction pathways</li> <li>Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>Transporters</li> <li>Metabolic enzymes</li> <li>Change in cell membrane structure and potential</li> <li>Kinases and phosphatases</li> <li>Transcription activators</li> <li>Change in chromatin structure and/or localized DNA topology</li> <li>Leaf production factors</li> </ul>
Down-regulated transcripts	<p>Responders to high FR:R light ratios</p> <p>Genes repressed by high FR:R light ratio</p> <p>Genes with discontinued expression or</p>	<ul style="list-style-type: none"> <li>Changes in pathways and processes operating in cells</li> <li>Changes in metabolisms other than far red stimulated pathways</li> </ul>	<ul style="list-style-type: none"> <li>Transcription factors</li> <li>Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> <li>Change in chromatin structure and/or DNA</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	unstable mRNA during high FR:R ratio light		topology <ul style="list-style-type: none"> <li>• Stability of factors for protein synthesis and degradation</li> <li>• Metabolic enzymes</li> <li>• Cell elongation factors</li> <li>• Flowering promotion factors</li> </ul>

#### Use of Promoters of Shade Avoidance Genes

Promoters of Shade Avoidance genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Shade Avoidance genes where the desired sequence is operably linked to a promoter of a Shade Avoidance gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoters are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.P. SULFUR RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Sulfur is one of the important macronutrients required by plants. It is taken up from the soil solution by roots as in the form of sulfate anion which higher plants are dependent on to fulfill their nutritional sulfur requirement. After uptake from the soil, sulfate is either accumulated and stored in vacuole or it is assimilated into various organic compounds, e.g. cysteine, glutathione, methionine, etc. Thus, plants also serve as nutritional sulfur sources for animals. Sulfur can be assimilated in one of two ways: it is either incorporated as sulfate in a reaction called sulfation, or it is first reduced to sulfide, the substrate for cysteine synthesis. In plants, majority of sulfur is assimilated in reduced form.

Sulfur comprises a small but vital fraction of the atoms in many protein molecules. As disulfide bridges, the sulfur atoms aid in stabilizing the folded proteins, such as cysteine residues. Cys is the first sulfur-containing amino acid, which in proteins forms disulfide bonds that may affect the tertiary structures and enzyme activities. This redox balance is mediated by the disulfide/thiol interchange of thioredoxin or glutaredoxin using NADPH as an electron donor. Sulfur can also become a thiol (SH) group participating in the active sites of some enzymes and some enzymes require the aid of small molecules that contain sulfur. In addition, the machinery of photosynthesis includes some sulfur-containing compounds, such as ferredoxin. Thus, sulfate assimilation plays important roles not only in sulfur nutrition but also in the ubiquitous process that may regulate the biochemical reactions of various metabolic pathways.

Deficiency of sulfur leads to a marked chlorosis in younger leaves, which may become white in color. Other symptoms of sulfur deficiency also include weak stems and reduced growth. Adding sulfur fertilizer to plants can increase root development and a deeper green color of the leaves in sulfur-deficient plants. However, sulfur is generally sufficient in soils for two reasons: it is a contaminant in potassium and other fertilizers and a product of industrial combustion. Sulfur limitation in plants is thus likely due to the limitation of the uptake and distribution of sulfate in plants. Seven cell type specific sulfate transporter genes have been isolated from Arabidopsis. In sulfate-starved plants, expression of the high-affinity transporter, AtST1-1, is induced in root epidermis and cortex for acquisition of sulfur. The low affinity transporter, AtST2-1 (AST68), accumulates in the root vascular tissue by sulfate starvation for root-to-shoot transport of sulfate. These studies have shown that the whole-plant process of sulfate transport is coordinately regulated by the expression of these 2 sulfate transporter genes.

under sulfur limited conditions. Recent studies have proposed that feeding of O-acetylserine, GSH and selenate may regulate the expression of AtST1-1 and AtST2-1 (AST68) in roots either positively or negatively. However, regulatory proteins that may directly control the expression of these genes have not been identified yet.

It has been established that there are regulatory interactions between assimilatory sulfate and nitrate reduction in plants. The two assimilatory pathways are very similar and well coordinated; deficiency for one element was shown to repress the other pathway. The coordination between them should be taken into consideration when one tries to alter one of pathways.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by simultaneously hybridizing two differentially labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) Science 270: 467-70). The Arabidopsis Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing 10,000 non-redundant ESTs, selected from 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases over the controls were identified, compared to the Ceres full-length cDNA and genomic sequence databanks, and identical Ceres clones identified. MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the results of this analysis, indicating those Ceres clones which are up or down regulated over controls, thereby indicating the Ceres clones which are sulfur response responsive genes.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Sulfur (relating to SMD 8034, SMD 8035)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A "-" is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Sulfur genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a "+" or "-" indication

Sulfur Genes Identified By Cluster Analyses Of Differential Expression

Sulfur Genes Identified By Correlation To Genes That Are Differentially

Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Sulfur genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Sulfur (relating to SMD 8034, SMD 8035) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

Sulfur Genes Identified By Correlation To Genes That Cause Physiological

Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Sulfur genes. A group in the MA\_clust is considered a Sulfur pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

Sulfur Genes Identified By Amino Acid Sequence Similarity

Sulfur genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Sulfur genes. Groups of Sulfur genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Sulfur pathway or network is a group of proteins that also exhibits Sulfur functions/utilities.

V.P.1. USE OF SULFUR RESPONSIVE GENES TO MODULATE  
PHENOTYPES

Sulfur responsive genes and gene products are useful to or modulate one or more of the following phenotypes:

- Growth
- Roots
- Stems
- Leaves
- Development
- Chloroplasts and Mitochondria
- Fruit Development
- Seed Development
- Seed storage proteins
- Senescence
- Differentiation
- Plastid/chloroplast and mitochondria differentiation
- Protection against oxidative damage
- regulation of enzymes via redox control by thiol groups
- Metabolic detoxification
- Photosynthesis
- Carbon dioxide fixation

To improve any of the phenotype(s) above, activities of one or more of the sulfur responsive genes or gene products can be modulated and tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed. As an example, a plant can be transformed according to Bechtold and Pelletier (1998, *Methods. Mol. Biol.* 82:259-266) and visually inspected for the desired phenotype or metabolically and/or functionally assayed according to Saito et al. (1994, *Plant Physiol.* 106: 887-95), Takahashi et al (1997, *Proc. Natl. Acad. Sci. USA* 94: 11102-07) and Koprivova et al. (2000, *Plant Physiol.* 122: 737-46).



V.P.2. USE OF SULFUR-RESPONSIVE GENES, GENE COMPONENTS AND  
PRODUCTS TO MODULATE BIOCHEMICAL ACTIVITIES

The activities of one or more of the sulfur responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below.

5 Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Growth , Differentiation and Development	<ul style="list-style-type: none"> <li>• Root</li> <li>• Leaf</li> <li>• Stem</li> <li>• Chloroplast/Mitochondria development/differentiation</li> <li>• Seed storage protein synthesis</li> </ul>	<p>Klein and Klein (1988) Mineral Nutrition, In CM Wilson and J Gregory, eds Fundamentals of Plant Science. Harper and Row Publishers, Inc., NY, p163</p> <p>Rost et al. (1984) The Absorption and Transport System, In R Bem, ed, Botany-A Brief Introduction to Plant Biology. John Wiley and Sons, NY, p96.</p> <p>Huluigie et al. (2000) Biochem Biophys Res Commun 271: 380-5</p> <p>Kapazoglou et al. (2000) Eur J Biochem 267: 352-60</p> <p>Kim et al. (1999) 209: 282-9</p>
Metabolisms	<ul style="list-style-type: none"> <li>• Sulfate uptake and transport</li> <li>• Cysteine Biosynthesis</li> <li>• Methionine biosynthesis</li> <li>• Carbon dioxide fixation in photosynthesis</li> </ul>	<p>Takahashi et al. (1997) Proc Natl Acad Sci USA 94: 11102-07</p> <p>Saito et al. (1992) Proc Natl Acad Sci USA 89: 8078-82</p> <p>Hesse et al. (1999) Amino Acids 16: 113-31</p> <p>Bourgis et al. (1999) Plant Cell 11: 1485-98</p> <p>Buchana (1991) Arch Biochem Biophys 288: 1-9</p>

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
	<ul style="list-style-type: none"> <li>Thioredoxin reduction</li> <li>Nitrogen metabolism</li> </ul>	<p>Leustek and Saito (1999) Plant Physiol 120: 637-43</p> <p>Mamedova et al. (1999) FEBS Lett 462: 421-4</p> <p>Koprivova et al. (2000) Plant Physiol. 122: 737-46</p> <p>Yamaguchi et al. (1999) Biosci Biotechnol Biochem 63: 762-6</p>
Plant Defenses	<ul style="list-style-type: none"> <li>Reduction of oxidative stress – oxygen metabolism and reactive oxygen species</li> <li>Detoxification of toxins, xenobiotics and heavy metals</li> <li>Defense against pathogens or microbes</li> <li>Disease prevention by secondary sulfur-containing compounds</li> <li>Activation of kinases and phosphatases</li> </ul>	<p>May et al. (1998) J Expt Bio 49: 649-67</p> <p>Kreuz et al. (1996) Plant Physiol 111: 349-53</p> <p>Zhao et al. (1998) Plant Cell 10: 359-70</p> <p>Kyung and Fleming (1997) J Food Prot 60: 67-71</p> <p>Fahey et al. (1997) Proc Natl Acad Sci USA 94: 10367-72</p> <p>Davis et al. (1999) Plant Cell 11: 1179-90</p>

Other biological activities that can be modulated by the sulfur responsive genes and gene products are listed in the REFERENCE Table. Assays for detecting such biological activities are described in the Protein Domain table.

- 5 Sulfur responsive genes are characteristically differentially transcribed in response to fluctuating sulfur levels or concentrations, whether internal or external to an organism or cell.

MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table reports the changes in transcript levels of various sulfur responsive genes.

Profiles of these different sulfur responsive genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
Up regulated transcripts	Responders to sulfur Application	<ul style="list-style-type: none"> <li>• Sulfur perception</li> <li>• Sulfur uptake and transport</li> <li>• Sulfur metabolism</li> <li>• Synthesis of secondary metabolites and/or proteins</li> <li>• Modulation of sulfur response transduction pathways</li> <li>• Specific gene transcription initiation</li> </ul>	<ul style="list-style-type: none"> <li>• Transporters</li> <li>• Metabolic enzymes</li> <li>• Change in cell membrane structure and potential</li> <li>• Kinases and phosphatases</li> <li>• Transcription activators</li> <li>• Change in chromatin structure and/or localized DNA topology</li> <li>• Redox control</li> </ul>
Down-regulated transcripts	<p>responder to sulfur repressors of sulfur state of metabolism</p> <p>Genes with discontinued expression or</p>	<ul style="list-style-type: none"> <li>• Negative regulation of sulfur pathways</li> <li>• Changes in pathways and processes</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Change in protein structure by phosphorylation (kinases) or dephosphorylation (phosphatases)</li> </ul>

TRANSCRIPT LEVELS	TYPE OF GENES	PHYSIOLOGICAL CONSEQUENCES	EXAMPLES OF BIOCHEMICAL ACTIVITY
	unstable mRNA in presence of sulfur	operating in cells <ul style="list-style-type: none"> <li>• Changes in other metabolisms than sulfur</li> </ul>	<ul style="list-style-type: none"> <li>• Change in chromatin structure and/or DNA topology</li> <li>• Stability of factors for protein synthesis and degradation</li> <li>• Metabolic enzymes</li> </ul>

#### Use of Promoters of Sulfur Responsive Genes

Promoters of Sulfur responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Sulfur responsive genes where the desired sequence is operably linked to a promoter of a Sulfur responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoters are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

V.Q. ZINC RESPONSIVE GENES, GENE COMPONENTS AND PRODUCTS

Phytoremediation of soils contaminated with toxic levels of heavy metals requires the understanding of plant metal transport and tolerance. The numerous *Arabidopsis thaliana* studies have given scientists the potential for dissection and elucidation of plant micronutrient/heavy metal uptake and accumulation pathways. It has been shown altered regulation of ZNT1, a Zn/Cd transporter, contributes to high Zn uptake. Isolation and characterization of Zn/Cd hyperaccumulation genes may allow expression in higher biomass plant species for efficient contaminated soil clean up. Identification of additional Zn transport, tolerance and nutrition-related genes involved in heavy metal accumulation will enable manipulation of increased uptake (for phytoremediation) as well as limitation of uptake or leak pathways that contribute to toxicity in crop plants. Additionally, Zn-binding ligands involved in Zn homeostasis or tolerance may be identified, as well as factors affecting the activity or expression of Zn binding transcription factors. Gene products acting in concert to effect Zn uptake, which would not have been identified in complementation experiments, including multimeric transporter proteins, could also be identified.

Microarray technology allows monitoring of gene expression levels for thousands of genes in a single experiment. This is achieved by simultaneously hybridizing two differentially labeled fluorescent cDNA pools to glass slides that contain spots of DNA (Schena et al. (1995) *Science* 270: 467-70). The *Arabidopsis* Functional Genomics Consortium (AFGC) has recently made public the results from such microarray experiments conducted with AFGC chips containing 10,000 non-redundant ESTs, selected from 37,000 randomly sequenced ESTs generated from mRNA of different tissues and developmental stages.

The sequences of the ESTs showing at least two-fold increases or decreases over the controls were identified, compared to the Ceres full-length cDNA and genomic sequence databanks, and identical Ceres clones identified. The Zn response information was then used in conjunction with the existing annotation to attribute biological function or utility to the full-length cDNA and corresponding genomic sequence.

The MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff Table(s) reports the transcript levels of the experiment (see EXPT ID: Zinc (relating to SMD 7310, SMD 7311)). For transcripts that had higher levels in the samples than the control, a "+" is shown. A

“-” is shown for when transcript levels were reduced in root tips as compared to the control. For more experimental detail see the Example section below.

Zinc genes are those sequences that showed differential expression as compared to controls, namely those sequences identified in the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff tables with a “+” or “-” indication

#### Zinc Genes Identified By Cluster Analyses Of Differential Expression

##### Zinc Genes Identified By Correlation To Genes That Are Differentially Expressed

As described above, the transcription profiles of genes that act together are well correlated. Applicants not only have identified the genes that are differentially expressed in the microarray experiments, but also have identified the genes that act in concert with them. The MA\_clust table indicates groups of genes that have well correlated transcription profiles and therefore participate in the same pathway or network.

A pathway or network of Zinc genes is any group in the MA\_clust that comprises a cDNA ID that also appears in Expt ID Zinc (relating to SMD 7310, SMD 7311) of the MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff table(s).

##### Zinc Genes Identified By Correlation To Genes That Cause Physiological Consequences

Additionally, the differential expression data and the phenotypic observations can be merged to identify pathways or networks of Zinc genes. A group in the MA\_clust is considered a Zinc pathway or network if the group comprises a cDNA ID that also appears in Knock-in or Knock-out tables that causes one or more of the phenotypes described in section above.

##### Zinc Genes Identified By Amino Acid Sequence Similarity

Zinc genes from other plant species typically encode polypeptides that share amino acid similarity to the sequences encoded by corn and Arabidopsis Zinc genes. Groups of Zinc genes are identified in the Protein Group table. In this table, any protein group that comprises a peptide ID that corresponds to a cDNA ID member of a Zinc pathway or network is a group of proteins that also exhibits Zinc functions/utilities.





V.Q.1. USE OF ZN TRANSPORT, TOLERANCE AND NUTRITION-  
RELATED GENES TO MODULATE PHENOTYPES

Changes in zinc concentration in the surrounding environment or in contact with a plant  
5 results in modulation of many genes and gene products. Examples of such zinc responsive genes  
and gene products are shown in the Reference, Sequence tables, Protein Group, Protein Group  
Matrix, MA\_diff and/or AFLP\_diff MA\_diff and/or AFLP\_diff and/or AFLP\_diff, and MA\_clust  
tables. These genes and/or products are responsible for effects on traits such as plant vigor and seed  
yield.

10 While zinc responsive polynucleotides and gene products can act alone, combinations of  
these polynucleotides also affect growth and development. Useful combinations include  
different zinc responsive polynucleotides and/or gene products that have similar transcription  
profiles or similar biological activities, and members of the same or similar biochemical  
15 pathways. In addition, the combination of a zinc responsive polynucleotide and/or gene product  
with another environmentally responsive polynucleotide is also useful because of the interactions  
that exist between hormone regulated pathways, stress pathways, nutritional pathways and  
development. Here, in addition to polynucleotides having similar transcription profiles and/or  
biological activities, useful combinations include polynucleotides that may have different  
transcription profiles but which participate in a common pathway.

20 Such zinc responsive genes and gene products can function to either increase or  
dampen the above phenotypes or activities either

- in response to changes in zinc concentration or
- in the absence of zinc fluctuations.

25 Zn transport, tolerance and nutrition-related genes and gene products can be used to  
alter or modulate one or more of the following phenotypes:

- Zn Uptake
- Transport of Zn or other heavy metals into roots
- Epidermal/cortical uptake
- Xylem loading
- Zn compartmentation
- Xylem unloading

- Phloem loading
- Efflux from cells to apoplast
- Sequestration in vacuoles/subcellular compartments.
- Zn tolerance
- Chelation of Zn
- Transport of Zn
- Metabolic and transcriptional control
- Activity of Zn binding enzymes
- Activity of Zn binding transcription factors

To improve any of the phenotype(s) above, activities of one or more of the Zn transport, tolerance and nutrition-related genes or gene products can be modulated and the plants can be tested by screening for the desired trait. Specifically, the gene, mRNA levels, or protein levels can be altered in a plant utilizing the procedures described herein and the phenotypes can be assayed, for example, in accordance to Lasat MM, Pence NS, Garvin DF, Ebbs SD, Kochian LV. J Exp Bot. 2000 Jan;51(342):71-9; Grotz N, Fox T, Connolly E, Park W, Gueriot ML, Eide D. Proc Natl Acad Sci U S A. 1998 Jun 9;95(12):7220-4; Crowder MW, Maiti MK, Banovic L, Makaroff CA. FEBS Lett. 1997 Dec 1;418(3):351-4; Hart JJ, Norvell WA, Welch RM, Sullivan LA, Kochian LV. Plant Physiol. 1998 Sep;118(1):219-26.

V.Q.2. USE OF ZN TRANSPORT, TOLERANCE AND NUTRITION-RELATED GENES TO MODULATE BIOCHEMICAL ACTIVITIES

Alternatively, the activities of one or more of the zinc responsive genes can be modulated to change biochemical or metabolic activities and/or pathways such as those noted below. Such biological activities can be measured according to the citations included in the Table below:

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Zn Uptake and Assimilation	Zn Influx	Lasat MM, Pence NS, Garvin DF, Ebbs SD, Kochian LV. J Exp Bot. 2000 Jan;51(342):71-9.
	Zn compartmentation	Hart JJ, Norvell WA, Welch RM, Sullivan LA, Kochian LV. Plant Physiol. 1998 Sep;118(1):219-26.
Zn binding by metabolic enzymes		Crowder MW, Maiti MK, Banovic L, Makaroff CA. FEBS Lett. 1997 Dec 1;418(3):351-4; Kenzior AL, Folk WR. FEBS Lett. 1998 Dec 4;440(3):425-9.
Zn binding by transcription factors		Newton AL, Sharpe BK, Kwan A, Mackay JP, Crossley M. J Biol Chem. 2000 May 19;275(20):15128- 34; Lopez Ribera I, Ruiz-Avila L, Puigdomenech P. Biochem Biophys Res Commun. 1997 Jul 18;236(2):510-6; de Pater S, Greco V, Pham K, Memelink J, Kijne J. Nucleic Acids Res. 1996 Dec 1;24(23):4624-31.

PROCESS	BIOCHEMICAL OR METABOLIC ACTIVITIES AND/OR PATHWAYS	CITATIONS INCLUDING ASSAYS
Synthesis of proteins to chelate Zn and other metals		Schafer HJ, Greiner S, Rausch T, Haag-Kerwer A. FEBS Lett. 1997 Mar 10;404(2-3):216-20.  Rausser WE. Cell Biochem Biophys. 1999;31(1):19-48.
Synthesis of metabolites to chelate Zn and other metals		Rausser WE. Cell Biochem Biophys. 1999;31(1):19-48.

Other biological activities that can be modulated by Zn transport, tolerance and nutrition-related genes and their products are listed in the Reference tables. Assays for detecting such biological activities are described in the Protein Domain table.

Zn transport, tolerance and nutrition-related genes are differentially transcribed in response to low Zn concentrations. The microarray comparison consists of probes prepared from root RNA of *A. thaliana* (Columbia) seedlings hydroponically grown in complete nutrient medium (control) and Zn deficient seedlings grown in -Zn nutrient medium (experimental). The data from this experiment reveal a number of types genes and gene products. MA\_diff and/or AFLP\_diff table reports the changes in transcript levels of various zinc responsive genes in entire seedlings at 1 and 6 hours after a plant was sprayed with a Hoagland's solution enriched with zinc as compared to seedlings sprayed with Hoagland's solution only.

The data from this time course can be used to identify a number of types of zinc responsive genes and gene products, including "early responding," "high zinc responders," "repressors of zinc deprivation pathways" and "zinc deprivation responders." Profiles of these different zinc responsive genes are shown in the Table below with examples of associated biological activities.

TRANSCRIPT LEVELS	TYPE OF GENE	PHYSIOLOGICAL CONSEQUENCE	EXAMPLES OF BIOCHEMICAL ACTIVITY
Upregulated transcripts (level at 1 hour $\approx$ 6 hours) (level at 1 hour > 6 hours)	Early responders to Zinc  Zinc Deprivation Responders	- Zinc Perception -Zinc Uptake - Modulation of Zinc Response Transduction Pathways - Specific Gene Transcription Initiation  -Repression of Pathways to Optimize Zinc Response Pathways	-Transcription Factors -Transporters  -Inhibit Transport of Zinc -Degradation
Level at 1 hour < 6 hours	Delayed Zinc Responders  Repressor of Zinc Deprivation Pathways	Negative Regulation of Zinc Pathways	-Zinc Metabolic Pathways
Down Regulated transcripts (Level at 1 hour $\approx$ 6 hours) (Level at 6 hours > 1 hour)	Early responder repressors of Zinc utilization Pathways	Negative Regulators of Zinc Utilization Pathways	Suppressing Zinc Requiring processes
Level at 1 hour > 6 hours	Genes with discontinued expression or unstable mRNA following Zinc uptake	Changes in pathways and processes operating in cells	

Use of Promoters of Zinc Responsive Genes

Promoters of Zinc responsive genes are useful for transcription of any desired polynucleotide or plant or non-plant origin. Further, any desired sequence can be transcribed in a similar temporal, tissue, or environmentally specific patterns as the Zinc responsive genes where the desired sequence is operably linked to a promoter of a Zinc responsive gene. The protein product of such a polynucleotide is usually synthesized in the same cells, in response to the same stimuli as the protein product of the gene from which the promoter was derived. Such promoter are also useful to produce antisense mRNAs to down-regulate the product of proteins, or to produce sense mRNAs to down-regulate mRNAs via sense suppression.

# VI. UTILITIES OF PARTICULAR INTEREST

Genes capable of modulating the phenotypes in the following table are useful produce the associated utilities in the table. Such genes can be identified by their cDNA ID number in the Knock-in and Knock-out tables.

5

<u>Phenotype Modulated by a Gene</u>	<u>Utilities</u>
Leaf shape	
<b>Cordate</b>	decrease wind opacity,
<b>Cup-shaped</b>	decrease lodging (plant fall over),
<b>Curled</b>	increase biomass by making larger or different shaped leaves,
<b>Lanceolate</b>	improve the efficiency of mechanical harvesting,
<b>Lobed</b>	decrease transpiration for better drought tolerance,
<b>Oval</b>	changing leaf shape to collect and absorb water,
<b>Ovate</b>	modulation of canopy structure and shading for altered irradiance close to the ground,
<b>Serrate</b>	enhanced uptake of pesticides (herbicides, fungicides, etc),
<b>Trident</b>	creation of ornamental leaf shapes,
<b>Undulate</b>	increase resistance to pathogens by decreasing amount of water that collects on leaves,
<b>Vertically Oblong</b>	change proportion of cell types in the leaves for enhanced photosynthesis,
<b>Other Shapes</b>	decreased transpiration, and enhanced accumulation of desirable compounds including secondary metabolites in specialized cells,
	decrease insect feeding,
<b>Long petioles</b>	decrease wind opacity,
<b>Short petioles</b>	decrease lodging (plant fall over), increase biomass by better positioning of the leaf blade,
	decrease insect feeding,
	decrease transpiration for better drought tolerance,
	position leaves most effectively for

**Phenotype**  
**Modulated by a**  
**Gene**

**Utilities**

		photosynthetic efficiency
	<b>Fused</b>	ornamental applications to make distinctive plants,
Reduced fertility	<b>Short siliques</b>	increase or decrease the number of seeds in a fruit, increasing fruit size, modulating fruit shape to better fit harvesting or packaging requirements, useful for controlling dehiscence and seed scatter
	<b>Reduced fertility Sterility</b>	useful in hybrid breeding programs, increasing fruit size, production of seedless fruit, useful as targets for gametocides, modulating fruit shape to better fit harvesting or packaging requirements, useful for controlling dehiscence and seed scatter
	<b>Flower size</b>	useful for edible flowers useful for flower derived products such as fragrances useful for modulating seed size and number in combination with seed-specific genes value in the ornamental industry
Stature	<b>Large</b> <b>Small</b>	increasing or decreasing plant biomass, optimizing plant stature to increase yield under various diverse environmental conditions, e.g., when water or nutrients are limiting,
	<b>Dwarfs</b>	decreasing lodging, increasing fruit number and size, controlling shading and canopy effects
Meristems		Change plant architecture,



**Phenotype  
Modulated by a  
Gene**

**Utilities**

increase or decrease number of leaves as well as change the types of leaves to increase biomass, improve photosynthetic efficiency, create new varieties of ornamental plants with enhanced leaf design, preventing flowering to optimize vegetative growth, control of apical dominance, increase or decrease flowering time to fit season, water or fertilizer schedules, change arrangement of leaves on the stem (phyllotaxy) to optimize plant density, decrease insect feeding, or decrease pathogen infection, increase number of trichome/glandular trichome producing leaves targets for herbicides, generate ectopic meristems and ectopic growth of vegetative and floral tissues and seeds and fruits

Stem

**Strong**

modify lignin content/composition for creation of harder woods or reduce difficulty/costs in pulping for

**Weak**

paper production or increase digestibility of forage crops, decrease lodging, modify cell wall polysaccharides in stems and fruits for improved texture and nutrition. increase biomass

**Late/Early Bolting**

Break the need for long vernalization of vernalization-dependent crops, e.g., winter wheat, thereby increasing yield decrease or increase generation time increase biomass

Lethals

**Embryo-lethal**

produce seedless fruit, use as herbicide targets

**Phenotype  
Modulated by a  
Gene**

**Utilities**

	<b>Embryo-defective</b>	produce seedless fruit, use as herbicide targets
	<b>Seedling</b>	use as herbicide targets, useful for metabolic engineering,
	<b>Pigment-lethals</b>	use as herbicide targets, increase photosynthetic efficiency
<b>Pigment</b>	<b>Dark Green</b>	Increase nutritional value, enhanced photosynthesis and carbon dioxide combustion and therefore increase plant vigor and biomass, enhanced photosynthetic efficiency and therefore increase plant vigor and biomass, prolong vegetative development, enhanced protection against pathogens,
	<b>YGV1</b>	Useful as targets for herbicides, increase photosynthetic efficiency and therefore increase plant vigor and biomass,
	<b>YGV2</b>	Useful as targets for herbicides, control of change from embryonic to adult organs, increase metabolic efficiency, increase photosynthetic efficiency and therefore increased plant vigor and biomass,
	<b>YGV3</b>	Useful as targets for herbicides, nitrogen sensing/uptake/usage, increase metabolic efficiency and therefore increased plant vigor and biomass,
	<b>Interveinal chlorosis</b>	to increase photosynthetic efficiency and therefore increase plant vigor and biomass to increase or decrease nitrogen transport and therefore increase plant vigor and

**Phenotype**  
**Modulated by a**  
**Gene**

**Utilities**

		biomass use as herbicide targets increase metabolic efficiency,
Roots	<b>Short (primary root)</b>	to access water from rainfall, to access rhizobia spray application, for anaerobic soils, useful to facilitate harvest of root crops,
	<b>Thick (primary root)</b>	useful for increasing biomass of root crops, for preventing plants dislodging during picking and harvesting, as root grafts, for animal feeds
	<b>Branching (primary root)</b>	modulation allows better access to water, minerals, fertilizers, rhizobia prevent soil erosion, increasing root biomass decrease root lodging,
	<b>Long (lateral roots)</b>	modulation allows improved access to water, nutrients, fertilizer, rhizobia, prevent soil erosion increase root biomass decrease root lodging modulation allows control on the depth of root growth in soil to access water and nutrients modulation allows hormonal control of root growth and development (size)
	<b>Agravitropic</b>	modulation allows control on the depth of root growth in soil
	<b>Curling (primary root)</b>	modulation allows hormonal control of root growth and development (size) useful in anaerobic soils in allowing roots to stay close to surface harvesting of root crops
	<b>Poor germination</b>	

**Phenotype**  
**Modulated by a**  
**Gene**

**Utilities**

Trichome	<b>Reduced Number</b> <b>Glabrous</b>	Genes useful for decreasing transpiration, increased production of glandular trichomes for oil or other secreted chemicals of value, use as deterrent for insect herbivory and oviposition
	<b>Increased Number</b>	modulation will increase resistance to UV light,
Wax mutants		decrease insect herbivory and oviposition, composition changes for the cosmetics industry, decrease transpiration, provide pathogen resistance, UV protection, modulation of leaf runoff properties and improved access for herbicides and fertilizers
Cotyledons		modulation of seeds structure in legumes, increase nutritional value, improve seedling competition under field conditions,
Seeds	<b>Transparent testa</b> <b>Light</b> <b>Dark</b>	genes useful for metabolic engineering anthocyanin and flavonoid pathways improved nutritional content
Flowers	<b>Other</b>	decrease petal abscission decrease pod shattering
Hypocotyl	<b>Long</b>  <b>Short</b>	to improve germination rates to improve plant survivability  to improve germination rates to improve plant survivability

## VII. ENHANCED FOODS

Animals require external supplies of amino acids that they cannot synthesize themselves.

5 Also, some amino acids are required in larger quantities. The nutritional values of plants for animals and humans can thus be modified by regulating the amounts of the constituent amino acids that occur as free amino acids or in proteins. For instance, higher levels of lysine and/or methionine would enhance the nutritional value of corn seed. Applicants herein provide several methods for modulating the amino acid content:

- 10 (1) expressing a naturally occurring protein that has a high percentage of the desired amino acid(s);
- (2) expressing a modified or synthetic coding sequence that has an enhanced percentage of the desired amino acids; or
- 15 (3) expressing the protein(s) that are capable of synthesizing more of the desired amino acids.

A specific example is expressing proteins with enhanced, for example, methionine content, preferentially in a corn or cereal seed used for animal nutrition or in the parts of plants used for nutritional purposes.

20 A protein is considered to have a high percentage of an amino acid if the amount of the desired amino acid is at least 1% of the total number of residues in a protein; more preferably 2% or greater. Amino acids of particular interest are tryptophan, lysine, methionine, phenylalanine, threonine leucine, valine, and isoleucine. Examples of naturally occurring proteins with a high percentage of any one of the amino acid of particular interest are listed in the Enhanced Amino Table.

25 The sequence(s) encoding the selected protein(s) are operably linked to a promoter and other regulatory sequences and transformed into a plant as described below. The promoter is chosen optimally for promoting the desired level of expression of the protein in the selected organ e.g. a promoter highly functional in seeds. Modifications may be made to the sequence encoding the protein to ensure protein transport into, for example, organelles or storage bodies or  
30 its accumulation in the organ. Such modifications may include addition of signal sequences at or near the N terminus and amino acid residues to modify protein stability or appropriate

5

VIII. USE OF NOVEL GENES TO FACILITATE EXPLOITATION OF PLANTS AS  
FACTORIES FOR THE SYNTHESIS OF VALUABLE MOLECULES

Plants and their constituent cells, tissues, and organs are factories that manufacture small organic molecules such as sugars, amino acids, fatty acids, vitamins, etc., as well as macromolecules such as proteins, nucleic acids, oils/fats and carbohydrates. Plants have long been a source of pharmaceutically beneficial chemicals; particularly, the secondary metabolites and hormone-related molecules synthesized by plants. Plants can also be used as factories to produce carbohydrates or lipids that comprises a carbon backbone useful as precursors of plastics, fiber, fuel, paper, pulp, rubber, solvents, lubricants, construction materials, detergents, and other cleaning materials. Plants can also generate other compounds that are of economic value, such as dyes, flavors, and fragrances. Both the intermediates as well as the end-products of plant bio-synthetic pathways have been found useful.

With the polynucleotides and polypeptides of the instant invention, modification of both in-vitro and in-vivo synthesis of such products is possible. One method of increasing the amount of either the intermediates or the end-products synthesized in a cell is to increase the expression of one or more proteins in the synthesis pathway as discussed below. Another method of increasing production of an intermediate is to inhibit expression of protein(s) that synthesize the end-product from the intermediate. Levels of end-products and intermediates can also be modified by changing the levels of enzymes that specifically change or degrade them. The kinds of molecules made can be also be modified by changing the genes encoding specific enzymes performing reactions at specific steps of the biosynthetic pathway. These genes can be from the same or a different organism. The molecular structures in the biosynthetic pathways can thus be modified or diverted into different branches of a pathway to make novel end-products.

Novel genes comprising selected promoters and sequences encoding enzymes are transformed into the selected plant to modify the levels, composition and/or structure of, without limitation:

- Terpenoids
- Alkaloids
- Hormones, including brassinosteroids
- Flavonoids

- Steroids
- Vitamins such as
  - Retinol
  - Riboflavin
  - Thiamine
- Caffeine
- Morphine and other alkaloids
- Peptides and amino acid synthesis
- Antioxidants
- Starches and lipids
- Fatty acids
- Fructose, mannose and other sugars
- Glycerolipid
- Citric acid
- Lignin
- Flavors
- Fragrances
- Essential oils
- Colors or dyes
- Gum
- Gels
- Waxes

The modifications are made by designing one or more novel genes per application comprising promoters, to ensure production of the enzyme(s) in the relevant cells, in the right amount, and polynucleotides encoding the relevant enzyme. The promoters and polynucleotides are the subject of this application. The novel genes are transformed into the relevant species using standard procedures. Their effects are measured by standard assays for the specific chemical/biochemical products.

These polynucleotides and proteins of the invention that participate in the relevant pathways and are useful for changing production of the above chemicals and biochemicals are



identified in the Reference tables by their enzyme function. More specifically, proteins of the invention that have the following enzymatic activities are of interest to modulate the corresponding pathways to produce precursors or final products noted above that are of industrial use. Biological activities of particular interest are listed below.

- 5           Other polynucleotides and proteins that regulate where, when and to what extent a pathway is active in a plant are extremely useful for modulating the synthesis and accumulation of valuable chemicals. These elements including transcription factors, proteins involved in signal transduction and other proteins in the control of gene expression are described elsewhere in this application.

Pathway Name	Enzyme Description	Comments
Alkaloid biosynthesis I	Morphine 6- dehydrogenase	Also acts on other alkaloids, including codeine, normorphine and ethylmorphine, but only very slowly on 7,8-saturated derivatives such as dihydromorphine and dihydrocodeine In the reverse direction, also reduces naloxone to the 6-alpha- hydroxy analog Activated by 2- mercaptoethanol
	Codeinone reductase (NADPH)	Stereospecifically catalyses the reversible reduction of codeinone to codeine, which is a direct precursor of morphine in the opium poppy plant, Papaver somniferum
	Salutaridine reductase (NADPH)	Stereospecifically catalyses the reversible reduction of salutaridine to salutaridinol, which is a direct precursor of morphinan alkaloids in the poppy plant, Papaver somniferum
	(S)-stylophine synthase	Catalyses an oxidative reaction that does not incorporate oxygen into the product Forms the second methylenedioxy bridge of the protoberberine alkaloid stylophine from oxidative ring closure of adjacent phenolic and methoxy groups of cheilanthifoline
	(S)-cheilanthifoline synthase	Catalyses an oxidative reaction that does not incorporate oxygen into the product Forms the methylenedioxy bridge of the protoberberine alkaloid cheilanthifoline from oxidative ring closure of adjacent phenolic and methoxy groups of scoulerine
	Salutaridine synthase	Forms the morphinan alkaloid salutaridine by intramolecular phenol oxidation of reticuline without the incorporation of oxygen into the product

	(S)-canadine synthase	Catalyses an oxidative reaction that does not incorporate oxygen into the product Oxidation of the methoxyphenol group of the alkaloid tetrahydrocolumbamine results in the formation of the methylenedioxy bridge of canadine
	Protopine 6-monooxygenase	Involved in benzophenanthridine alkaloid synthesis in higher plants
	Dihydrosanguinarine 10-monooxygenase	Involved in benzophenanthridine alkaloid synthesis in higher plants
	Monophenol monooxygenase	A group of copper proteins that also catalyse the reaction of EC 1.10.3.1, if only 1,2-benzenediols are available as substrate
	L-amino acid oxidase	
	1,2-dehydroreticuline reductase (NADPH)	Stereospecifically reduces the 1,2-dehydroreticuline ion to (R)-reticuline, which is a direct precursor of morphinan alkaloids in the poppy plant, papaver somniferum The enzyme does not catalyse the reverse reaction to any significant extent under physiological conditions
	Dihydrobenzophenanthridine oxidase	Also catalyzes: dihydrochelirubine + O(2) = chelirubine + H(2)O(2) Also catalyzes: dihydromacarpine + O(2) = macarpine + H(2)O(2) Found in higher plants Produces oxidized forms of the benzophenanthridine alkaloids
	Reticuline oxidase	The product of the reaction, (S)-scoulerine, is a precursor of protopine, protoberberine and benzophenanthridine alkaloid biosynthesis in plants Acts on (S)-reticuline and related compounds, converting the N- methyl group into the methylene bridge (berberine bridge[PRIME]) of (S)-tetrahydroprotoberberines

	3[PRIME]-hydroxy-N-methyl-(S)-coclaaurine 4[PRIME]-O-methyltransferase	Involved in isoquinoline alkaloid metabolism in plants Has also been shown to catalyse the methylation of (R,S)- laudanosoline, (S)-3[PRIME]-hydroxycoclaaurine and (R,S)-7-O-methylnorlaudanosoline
	(S)-scoulerine 9-O-methyltransferase	The product of this reaction is a precursor for protoberberine alkaloids in plants
	Columbamine O-methyltransferase	The product of this reaction is a protoberberine alkaloid that is widely distributed in the plant kingdom Distinct in specificity from EC 2.1.1.88
	10-hydroxydihydrosan guinarine 10-O-methyltransferase	Part of the pathway for synthesis of benzophenanthridine alkaloids in plants
	12-hydroxydihydrochelirubine 12-O-methyltransferase	Part of the pathway for synthesis of benzophenanthridine alkaloid macarpine in plants
	(R,S)-norcoclaaurine 6-O-methyltransferase	Norcoclaaurine is 6,7-dihydroxy-1-[(4-hydroxyphenyl)methyl]-1,2,3,4-tetrahydroisoquinoline The enzyme will also catalyse the 6-O-methylation of (R,S)- norlaudanosoline to form 6-O-methyl-norlaudanosoline, but this alkaloid has not been found to occur in plants
	Salutaridinol 7-O-acetyltransferase	At higher pH values the product, 7-O-acetylsalutaridinol, spontaneously closes the 4->5 oxide bridge by allylic elimination to form the morphine precursor thebaine From the opium poppy plant, Papaver somniferum
	Aspartate aminotransferase	Also acts on L-tyrosine, L-phenylalanine and L-tryptophan. This activity can be formed from EC 2.6.1.57 by controlled proteolysis
	Tyrosine aminotransferase	L-phenylalanine can act instead of L-tyrosine The mitochondrial enzyme may be identical with EC 2.6.1.1 The three isoenzymic forms are interconverted by EC 3.4.22.4

	Aromatic amino acid transferase	L-methionine can also act as donor, more slowly Oxaloacetate can act as acceptor Controlled proteolysis converts the enzyme to EC 2.6.1.1
	Tyrosine decarboxylase	The bacterial enzyme also acts on 3-hydroxytyrosine and, more slowly, on 3-hydroxyphenylalanine
	Aromatic-L-amino-acid decarboxylase	Also acts on L-tryptophan, 5-hydroxy-L-tryptophan and dihydroxy- L-phenylalanine (DOPA)
Alkaloid biosynthesis II	Tropine dehydrogenase	Oxidizes other tropan-3-alpha-ols, but not the corresponding beta-derivatives
	Tropinone reductase	
	Hyoscyamine (6S)-dioxygenase	
	6-beta-hydroxyhyoscyamine epoxidase	
	Amine oxidase (copper-containing)	A group of enzymes including those oxidizing primary amines, diamines and histamine One form of EC 1.3.1.15 from rat kidney also catalyses this reaction
	Putrescine N-methyltransferase	
	Ornithine decarboxylase	
	Oxalyl-CoA decarboxylase	
	Phenylalanine ammonia-lyase	May also act on L-tyrosine
Androgen and estrogen metabolism	3-beta-hydroxy-delta(5)-steroid dehydrogenase	Acts on 3-beta-hydroxyandrost-5-en-17-one to form androst-4-ene- 3,17-dione and on 3-beta-hydroxypregn-5-en-20-one to form progesterone
	11-beta-hydroxysteroid dehydrogenase	
	Estradiol 17-alpha-dehydrogenase	
	3-alpha-hydroxy-5-beta-androstane-17-one 3-alpha-	

	dehydrogenase	
	3-alpha (17-beta)-hydroxysteroid dehydrogenase (NAD+)	Also acts on other 17-beta-hydroxysteroids, on the 3-alpha-hydroxy group of pregnanes and bile acids, and on benzene dihydrodiol Different from EC 1.1.1.50 or EC 1.1.1.213
	3-alpha-hydroxysteroid dehydrogenase (B-specific)	Acts on other 3-alpha-hydroxysteroids and on 9-, 11- and 15-hydroxyprostaglandin B-specific with respect to NAD(+) or NADP(+) (cf. EC 1.1.1.213)
	3(or 17)beta-hydroxysteroid dehydrogenase	Also acts on other 3-beta- or 17-beta-hydroxysteroids (cf EC 1.1.1.209)
	Estradiol 17 beta-dehydrogenase	Also acts on (S)-20-hydroxypregn-4-en-3-one and related compounds, oxidizing the (S)-20-group B-specific with respect to NAD(P)(+)
	Testosterone 17-beta-dehydrogenase	
	Testosterone 17-beta-dehydrogenase (NADP+)	Also oxidizes 3-hydroxyhexobarbital to 3-oxohexobarbital
	Steroid 11-beta-monooxygenase	Also hydroxylates steroids at the 18-position, and converts 18-hydroxycorticosterone into aldosterone
	Estradiol 6-beta-monooxygenase	
	Androst-4-ene-3,17-dione monooxygenase	Has a wide specificity A single enzyme from <i>Cylindrocarpum radicola</i> (EC 1.14.13.54) catalyses both this reaction and that catalysed by EC 1.14.99.4
	3-oxo-5-alpha-steroid 4-dehydrogenase	
	3-oxo-5-beta-steroid 4-	

	dehydrogenase	
	UDP-glucuronosyltransferase	Family of enzymes accepting a wide range of substrates, including phenols, alcohols, amines and fatty acids Some of the activities catalysed were previously listed separately as EC 2.4.1.42, EC 2.4.1.59, EC 2.4.1.61, EC 2.4.1.76, EC 2.4.1.77, EC 2.4.1.84, EC 2.4.1.107 and EC 2.4.1.108 A temporary nomenclature for the various forms whose delineation is in a state of flux
	Steroid sulfotransferase	Broad specificity resembling EC 2.8.2.2, but also acts on estrone
	Alcohol sulfotransferase	Primary and secondary alcohols, including aliphatic alcohols, ascorbate, chloramphenicol, ephedrine and hydroxysteroids, but not phenolic steroids, can act as acceptors (cf. EC 2.8.2.15)
	Estrone sulfotransferase	
	Arylsulfatase	A group of enzymes with rather similar specificities
	Steryl-sulfatase	Also acts on some related steryl sulfates
	17-alpha-hydroxyprogesterone aldolase	
	Steroid delta-isomerase	
C21-Steroid hormone metabolism	3-beta-hydroxy-delta(5)-steroid dehydrogenase	Acts on 3-beta-hydroxyandrost-5-en-17-one to form androst-4-ene-3,17-dione and on 3-beta-hydroxypregn-5-en-20-one to form progesterone
	11-beta-hydroxysteroid dehydrogenase	
	20-alpha-hydroxysteroid dehydrogenase	A-specific with respect to NAD(P)(+)
	3-alpha-hydroxysteroid dehydrogenase (B-	Acts on other 3-alpha-hydroxysteroids and on 9-, 11- and 15-hydroxyprostaglandin B-specific with

	specific)	respect to NAD(+) or NADP(+) (cf. EC 1.1.1.213)
	3-alpha(or 20-beta)-hydroxysteroid dehydrogenase	The 3-alpha-hydroxyl group or 20-beta-hydroxyl group of pregnane and androstane steroids can act as donors
	Steroid 11-beta-monooxygenase	Also hydroxylates steroids at the 18-position, and converts 18-hydroxycorticosterone into aldosterone
	Corticosterone 18-monooxygenase	
	Cholesterol monooxygenase (side-chain cleaving)	The reaction proceeds in three stages, with hydroxylation at C-20 and C-22 preceding scission of the side-chain at C-20
	Steroid 21-monooxygenase	
	Progesterone 11-alpha-monooxygenase	
	Steroid 17-alpha-monooxygenase	
	Cholestenone 5-beta-reductase	
	Cortisone beta-reductase	
	Progesterone 5-alpha-reductase	Testosterone and 20-alpha-hydroxy-4-pregnen-3-one can act in place of progesterone
	3-oxo-5-beta-steroid 4-dehydrogenase	
	Steroid delta-isomerase	
Flavonoids, stilbene and lignin biosynthesis	Coniferyl-alcohol dehydrogenase	Specific for coniferyl alcohol; does not act on cinnamyl alcohol, 4-coumaryl alcohol or sinapyl alcohol
	Cinnamyl-alcohol dehydrogenase	Acts on coniferyl alcohol, sinapyl alcohol, 4-coumaryl alcohol and



		cinnamyl alcohol (cf. EC 1.1.1.194)
	Dihydrokaempferol 4-reductase	Also acts, in the reverse direction, on (+)-dihydroquercetin and (+)-dihydromyricetin Each dihydroflavonol is reduced to the corresponding cis-flavon- 3,4-diol NAD(+) can act instead of NADP(+), more slowly Involved in the biosynthesis of anthocyanidins in plants
	Flavonone 4-reductase	Involved in the biosynthesis of 3-deoxyanthocyanidins from flavonones such as naringenin or eriodictyol
	Peroxidase	
	Caffeate 3,4-dioxygenase	
	Naringenin 3-dioxygenase	
	Trans-cinnamate 4-monooxygenase	Also acts on NADH, more slowly
	Trans-cinnamate 2-monooxygenase	
	Flavonoid 3[PRIME]-monooxygenase	Acts on a number of flavonoids, including naringenin and dihydrokaempferol Does not act on 4-coumarate or 4-coumaroyl-CoA
	Monophenol monooxygenase	A group of copper proteins that also catalyse the reaction of EC 1.10.3.1, if only 1,2-benzenediols are available as substrate
	Cinnamoyl-CoA reductase	Also acts on a number of substituted cinnamoyl esters of coenzyme A
	Caffeoyl-CoA O-methyltransferase	
	Luteolin O-methyltransferase	Also acts on luteolin-7-O-beta-D-glucoside
	Caffeate O-methyltransferase	3,4-dihydroxybenzaldehyde and catechol can act as acceptor, more

		slowly
	Apigenin 4[PRIME]-O- methyltransferase	Converts apigenin into acacetin Naringenin (5,7,4[PRIME]- trihydroxyflavonone) can also act as acceptor, more slowly
	Quercetin 3-O- methyltransferase	Specific for quercetin. Related enzymes bring about the 3-O- methylation of other flavonols, such as galangin and kaempferol
	Isoflavone-7-O- beta-glucoside 6[PRIME][PRIME] -O- malonyltransferase	The 6-position of the glucose residue of formononetin can also act as acceptor Some other 7-O-glucosides of isoflavones, flavones and flavonols can also act, more slowly
	Pinosylvin synthase	Not identical with EC 2.3.1.74 or EC 2.3.1.95
	Naringenin- chalcone synthase	In the presence of NADH and a reductase, 6[PRIME]-deoxychalcone is produced
	Trihydroxystilbene synthase	Not identical with EC 2.3.1.74 or EC 2.3.1.146
	Quinate O- hydroxycinnamoyl transferase	Caffeoyl-CoA and 4-coumaroyl-CoA can also act as donors, more slowly Involved in the biosynthesis of chlorogenic acid in sweet potato and, with EC 2.3.1.98 in the formation of caffeoyl-CoA in tomato
	Coniferyl-alcohol glucosyltransferase	Sinapyl alcohol can also act as acceptor
	2-coumarate O- beta- glucosyltransferase	Coumarinate (cis-2- hydroxycinnamate) does not act as acceptor
	Scopoletin glucosyltransferase	
	Flavonol-3-O- glucoside L- rhamnosyltransferase	Converts flavonol 3-O-glucosides to 3-O-rutinosides Also acts, more slowly, on rutin, quercetin 3-O- galactoside and flavonol O3- rhamnosides
	Flavone 7-O-beta- glucosyltransferase	A number of flavones, flavonones and flavonols can function as acceptors Different from EC 2.4.1.91
	Flavonol 3-O- glucosyltransferase	Acts on a variety of flavonols, including quercetin and quercetin 7-O-

		glucoside Different from EC 2.4.1.81
	Flavone apiosyltransferase	7-O-beta-D-glucosides of a number of flavonoids and of 4-substituted phenols can act as acceptors
	Coniferin beta- glucosidase	Also hydrolyzes syringin, 4-cinnamyl alcohol beta-glucoside, and, more slowly, some other aryl beta- glycosides A plant cell-wall enzyme involved in the biosynthesis of lignin
	Beta-glucosidase	Wide specificity for beta-D- glucosides. Some examples also hydrolyse one or more of the following: beta-D-galactosides, alpha- L- arabinosides, beta-D-xylosides, and beta-D-fucosides
	Chalcone isomerase	
	4-coumarate--CoA ligase	

Pathway Name	Enzyme Description	Enzyme Comments
Ascorbate and aldarate metabolism	D-threo-aldose 1-dehydrogenase	Acts on L-fucose, D-arabinose and L-xylose The animal enzyme was also shown to act on L-arabinose, and the enzyme from <i>Pseudomonas caryophylli</i> on L-glucose
	L-threonate 3-dehydrogenase	
	Glucuronate reductase	Also reduces D-galacturonate May be identical with EC 1.1.1.2
	Glucuronolactone reductase	
	L-arabinose 1-dehydrogenase	
	L-galactonolactone oxidase	Acts on the 1,4-lactones of L-galactonic, D-galactonic, L-fuconic, D-arabinic and D-threonic acids Not identical with EC 1.1.3.8 (cf. EC 1.3.2.3)
	L-gulonolactone oxidase	The product spontaneously isomerizes to L-ascorbate
	L-ascorbate oxidase	
	L-ascorbate peroxidase	
	Ascorbate 2,3-dioxygenase	
	2,5-dioxovalerate dehydrogenase	

	Aldehyde dehydrogenase (NAD+)	Wide specificity, including oxidation of D-glucuronolactone to D-glucarate
	Galactonolactone dehydrogenase	Cf. EC 1.1.3.24
	Monodehydroascorbate reductase (NADH)	
	Glutathione dehydrogenase (ascorbate)	
	L-arabinonolactonase	
	Gluconolactonase	Acts on a wide range of hexono-1,5-lactones
	Uronolactonase	
	1,4-lactonase	Specific for 1,4-lactones with 4-8 carbon atoms Does not hydrolyse simple aliphatic esters, acetylcholine, sugar lactones or substituted aliphatic lactones, e.g. 3-hydroxy-4-butyrolactone
	2-dehydro-3-deoxyglucarate aldolase	
	L-arabinonate dehydratase	
	Glucarate dehydratase	
	5-dehydro-4-deoxyglucarate dehydratase	
	Galactarate dehydratase	
	2-dehydro-3-	

	deoxy-L-arabinonate dehydratase	
Carbon fixation	Malate dehydrogenase	Also oxidizes some other 2-hydroxydicarboxylic acids
	Malate dehydrogenase (decarboxylating)	Does not decarboxylates added oxaloacetate
	Malate dehydrogenase (oxaloacetate decarboxylating) (NADP+)	Also decarboxylates added oxaloacetate
	Malate dehydrogenase (NADP+)	Activated by light
	Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating)	
	Transketolase	Wide specificity for both reactants, e.g. converts hydroxypyruvate and R-CHO into CO(2) and R-CHOH-CO-CH(2)OH Transketolase from <i>Alcaligenes faecalis</i> shows high activity with D-erythrose as acceptor
	Aspartate aminotransferase	Also acts on L-tyrosine, L-phenylalanine and L-tryptophan. This activity can be formed from EC 2.6.1.57 by controlled

		proteolysis
	Alanine aminotransferase	2-aminobutanoate acts slowly instead of alanine
	Sedoheptulokinase	
	Phosphoribulokinase	
	Pyruvate kinase	UTP, GTP, CTP, ITP and dATP can also act as donors Also phosphorylates hydroxylamine and fluoride in the presence of CO(2)
	Phosphoglycerate kinase	
	Pyruvate, phosphate dikinase	
	Fructose-bisphosphatase	The animal enzyme also acts on sedoheptulose 1,7- bisphosphate
	Sedoheptulose-bisphosphatase	
	Phosphoenolpyruvate carboxylase	
	Ribulose-bisphosphate carboxylase	Will utilize O(2) instead of CO(2), forming 3-phospho- D-glycerate and 2- phosphoglycolate
	Phosphoenolpyruvate carboxykinase (ATP)	
	Fructose-bisphosphate aldolase	Also acts on (3S,4R)-ketose 1- phosphates The yeast and bacterial enzymes are zinc proteins The enzymes increase electron-attraction by the carbonyl

		group, some (Class I) forming a protonated imine with it, others (Class II), mainly of microbial origin, polarizing it with a metal ion, e.g zinc
	Phosphoketolase	
	Ribulose-phosphate 3-epimerase	Also converts D-erythrose 4-phosphate into D-erythrulose 4-phosphate and D-threose 4-phosphate
	Triosephosphate isomerase	
	Ribose 5-phosphate epimerase	Also acts on D-ribose 5-diphosphate and D-ribose 5-triphosphate
Phenylalanine metabolism	(R)-4-hydroxyphenyllactate dehydrogenase	Also acts, more slowly, on (R)-3-phenyllactate, (R)-3-(indole-3-yl)lactate and (R)-lactate
	Hydroxyphenylpyruvate reductase	Also acts on 3-(3,4-dihydroxyphenyl)lactate Involved with EC 2.3.1.140 in the biosynthesis of rosmarinic acid
	Aryl-alcohol dehydrogenase	A group of enzymes with broad specificity towards primary alcohols with an aromatic or cyclohex-1-ene ring, but with low or no activity towards short-chain aliphatic alcohols
	Peroxidase	
	Catechol 1,2-	Involved in the

[illegible]



	dioxygenase	metabolism of nitro-aromatic compounds by a strain of <i>Pseudomonas putida</i>
	2,3-dihydroxybenzoate 3,4-dioxygenase	
	3-carboxyethylcatechol 2,3-dioxygenase	
	Catechol 2,3-dioxygenase	The enzyme from <i>Alcaligenes</i> sp. strain O-1 has also been shown to catalyse the reaction: 3-Sulfocatechol + O(2) + H(2)O = 2-hydroxymuconate + bisulfite. It has been referred to as 3-sulfocatechol 2,3-dioxygenase. Further work will be necessary to show whether or not this is a distinct enzyme
	4-hydroxyphenylpyruvate dioxygenase	
	Protocatechuate 3,4-dioxygenase	
	Hydroxyquinol 1,2-dioxygenase	The product isomerizes to 2-maleylacetate (cis-hex-2-enedioate) Highly specific; catechol and pyrogallol are acted on at less than 1% of the rate at which benzene-1,2,4-triol

		is oxidized
	Protocatechuate 4,5-dioxygenase	
	Phenylalanine 2-monooxygenase	Also catalyses a reaction similar to that of EC 1.4.3.2, forming 3-phenylpyruvate, NH(3) and H(2)O(2), but more slowly
	Anthranilate 1,2-dioxygenase (deaminating, decarboxylating)	
	Benzoate 1,2-dioxygenase	A system, containing a reductase which is an iron-sulfur flavoprotein (FAD), and an iron-sulfur oxygenase
	Toluene dioxygenase	A system, containing a reductase which is an iron-sulfur flavoprotein (FAD), an iron-sulfur oxygenase, and a ferredoxin. Some other aromatic compounds, including ethylbenzene, 4-xylene and some halogenated toluenes, are converted into the corresponding cis-dihydrodiols
	Naphthalene 1,2-dioxygenase	A system, containing a reductase which is an iron-sulfur flavoprotein (FAD), an iron-sulfur oxygenase, and ferredoxin
	Benzene 1,2-dioxygenase	A system, containing a reductase which is

		an iron-sulfur flavoprotein, an iron-sulfur oxygenase and ferredoxin
	Salicylate 1-monooxygenase	
	Trans-cinnamate 4-monooxygenase	Also acts on NADH, more slowly
	Benzoate 4-monooxygenase	
	4-hydroxybenzoate 3-monooxygenase	Most enzymes from <i>Pseudomonas</i> are highly specific for NAD(P)H (cf EC 1.14.13.33)
	3-hydroxybenzoate 4-monooxygenase	Also acts on a number of analogs of 3-hydroxybenzoate substituted in the 2, 4, 5 and 6 positions
	3-hydroxybenzoate 6-monooxygenase	Also acts on a number of analogs of 3-hydroxybenzoate substituted in the 2, 4, 5 and 6 positions NADPH can act instead of NADH, more slowly
	4-hydroxybenzoate 3-monooxygenase (NAD(P)H)	The enzyme from <i>Corynebacterium cyclohexanicum</i> is highly specific for 4-hydroxybenzoate, but uses NADH and NADPH at approximately equal rates (cf. EC 1.14.13.2). It is less specific for NADPH than EC 1.14.13.2
	Anthranilate 3-monooxygenase (deaminating)	The enzyme from <i>Aspergillus niger</i> is an iron protein; that

624

	(NADP+)	
	Coumarate reductase	
	Cis-1,2-dihydrobenzene-1,2-diol dehydrogenase	
	Cis-1,2-dihydro-1,2-dihydroxynaphthalene dehydrogenase	Also acts, at half the rate, on cis-anthracene dihydrodiol and cis-phenanthrene dihydrodiol
	2-enoate reductase	Acts, in the reverse direction, on a wide range of alkyl and aryl alpha,beta-unsaturated carboxylate ions 2-butenate was the best substrate tested
	Maleylacetate reductase	
	Phenylalanine dehydrogenase	The enzyme from <i>Bacillus badius</i> and <i>Sporosarcina ureae</i> are highly specific for L-phenylalanine, that from <i>Bacillus sphaericus</i> also acts on L-tyrosine
	L-amino acid oxidase	
	Amine oxidase (flavin-containing)	Acts on primary amines, and usually also on secondary and tertiary amines
	Amine oxidase (copper-containing)	A group of enzymes including those oxidizing primary amines, diamines and histamine One form of EC 1.3.1.15 from rat kidney also catalyses this

		reaction
	D-amino-acid dehydrogenase	Acts to some extent on all D-amino acids except D-aspartate and D-glutamate
	Aralkylamine dehydrogenase	Phenazine methosulfate can act as acceptor Acts on aromatic amines and, more slowly, on some long-chain aliphatic amines, but not on methylamine or ethylamine (cf EC 1.4.99.3)
	Glutamine N-phenylacetyltransferase	
	Acetyl-CoA C-acyltransferase	
	D-amino-acid N-acetyltransferase	
	Phenylalanine N-acetyltransferase	Also acts, more slowly, on L-histidine and L-alanine
	Glycine N-benzoyltransferase	Not identical with EC 2.3.1.13 or EC 2.3.1.68
	Aspartate aminotransferase	Also acts on L-tyrosine, L-phenylalanine and L-tryptophan. This activity can be formed from EC 2.6.1.57 by controlled proteolysis
	D-alanine aminotransferase	Acts on the D-isomers of leucine, aspartate, glutamate, aminobutyrate, norvaline and asparagine

	Tyrosine aminotransferase	L-phenylalanine can act instead of L-tyrosine The mitochondrial enzyme may be identical with EC 2.6.1.1 The three isoenzymic forms are interconverted by EC 3.4.22.4
	Aromatic amino acid transferase	L-methionine can also act as donor, more slowly Oxaloacetate can act as acceptor Controlled proteolysis converts the enzyme to EC 2.6.1.1
	Histidinol-phosphate aminotransferase	
	3-oxoadipate CoA-transferase	
	3-oxoadipate enol-lactonase	Acts on the product of EC 4.1.1.44
	Carboxymethylenbutenolide	
	2-pyrone-4,6-dicarboxylate lactonase	The product isomerizes to 4-oxalomesaconate
	Hippurate hydrolase	Acts on various N-benzoylamino acids
	Amidase	
	Acylphosphatase	
	2-hydroxymuconate-semialdehyde hydrolase	
	Aromatic-L-amino-acid decarboxylase	Also acts on L-tryptophan, 5-hydroxy-L-tryptophan and dihydroxy-L-

		phenylalanine (DOPA)
	Phenylpyruvate decarboxylase	Also acts on indole-3-pyruvate
	4-carboxymuconolactone decarboxylase	
	O-pyrocatechuate decarboxylase	
	Phenylalanine decarboxylase	Also acts on tyrosine and other aromatic amino acids
	4-hydroxybenzoate decarboxylase	
	Protocatechuate decarboxylase	
	Benzoylformate decarboxylase	
	4-oxalocrotonate decarboxylase	Involved in the meta-cleavage pathway for the degradation of phenols, cresols and catechols
	4-hydroxy-4-methyl-2-oxoglutarate aldolase	Also acts on 4-hydroxy-4-methyl-2-oxoadipate and 4-carboxy-4-hydroxy-2-oxohexadioate
	2-oxopent-4-enoate hydratase	Also acts, more slowly, on cis-2-oxohex-4-enoate, but not on the trans-isomer
	Phenylalanine ammonia-lyase	May also act on L-tyrosine
	Phenylalanine racemase (ATP-hydrolysing)	
	Mandelate racemase	
	Phenylpyruvate	Also acts on other



	tautomerase	arylpyruvates
	5-carboxymethyl-2-hydroxymuconate delta-isomerase	
	Muconolactone delta-isomerase	
	Muconate cycloisomerase	Also acts, in the reverse reaction, on 3-methyl-cis-cis-hexa-dienedioate and, very slowly, on cis-trans-hexadienedioate Not identical with EC 5.5.1.7 or EC 5.5.1.11
	3-carboxy-cis,cis-muconate cycloisomerase	
	Carboxy-cis,cis-muconate cyclase	
	Chloromuconate cycloisomerase	Spontaneous elimination of HCl produces cis-4-carboxymethylenebut-2-en-4-olide Also acts in reverse direction on 2-chloro-cis,cis-muconate Not identical with EC 5.5.1.1 or EC 5.5.1.11
	Phenylacetate--CoA ligase	Phenoxyacetate can replace phenylacetate
	Benzoate--CoA ligase	Also acts on 2-, 3- and 4-fluorobenzoate, but only very slowly on the corresponding chlorobenzoates

	4-hydroxybenzoate--CoA ligase	
	Phenylacetate--CoA ligase	Also acts, more slowly, on acetate, propanoate and butanoate, but not on hydroxy derivatives of phenylacetate and related compounds
Phenylalanine, tyrosine and tryptophan biosynthesis	Quinate 5-dehydrogenase	
	Shikimate 5-dehydrogenase	
	Quinate dehydrogenase (pyrroloquinoline-quinone)	
	Phenylalanine 4-monoxygenase	
	Prephenate dehydrogenase	This enzyme in the enteric bacteria also possesses chorismate mutase activity (EC 5.4.99.5) and converts chorismate into prephenate
	Prephenate dehydrogenase (NADP+)	
	Cyclohexadienyl dehydrogenase	Also acts on prephenate and D-prephenyllactate (cf. EC 1.3.1.12)
	2-methyl-branched-chain-enoyl-CoA reductase	From <i>Ascaris suum</i> The reaction proceeds only in the presence of another flavoprotein (ETF=[PRIME]Electron-Transferring Flavoprotein[PRIME])
	Phenylalanine	The enzyme from

	dehydrogenase	Bacillus badius and Sporosarcina ureae are highly specific for L-phenylalanine, that from Bacillus sphaericus also acts on L-tyrosine
	L-amino acid oxidase	
	Anthranilate phosphoribosyltransferase	In some organisms, this enzyme is part of a multifunctional protein together with one or more components of the system for biosynthesis of tryptophan (EC 4.1.1.48, EC 4.1.3.27, EC 4.2.1.20, and EC 5.3.1.24)
	3-phosphoshikimate 1-carboxyvinyltransferase	
	Aspartate aminotransferase	Also acts on L-tyrosine, L-phenylalanine and L-tryptophan. This activity can be formed from EC 2.6.1.57 by controlled proteolysis
	Tyrosine aminotransferase	L-phenylalanine can act instead of L-tyrosine The mitochondrial enzyme may be identical with EC 2.6.1.1 The three isoenzymic forms are interconverted by EC 3.4.22.4

	Aromatic amino acid transferase	L-methionine can also act as donor, more slowly Oxaloacetate can act as acceptor Controlled proteolysis converts the enzyme to EC 2.6.1.1
	Histidinol-phosphate aminotransferase	
	Shikimate kinase	
	Indole-3-glycerol-phosphate synthase	In some organisms, this enzyme is part of a multifunctional protein together with one or more components of the system for biosynthesis of tryptophan (EC 2.4.2.18, EC 4.1.3.27, EC 4.2.1.20, and EC 5.3.1.24)
	2-dehydro-3-deoxyphosphonate aldolase	
	Anthranilate synthase	In some organisms, this enzyme is part of a multifunctional protein together with one or more components of the system for biosynthesis of tryptophan (EC 2.4.2.18, EC 4.1.1.48, EC 4.2.1.20, and EC 5.3.1.24) The native enzyme in the complex with uses

# THE

	3-dehydroquinase synthase	The hydrogen atoms on C-7 of the substrate are retained on C-2 of the products
	Chorismate synthase	Shikimate is numbered so that the double-bond is between C-1 and C-2, but some earlier papers numbered in the reverse direction
	Phosphoribosylanthranilate isomerase	In some organisms, this enzyme is part of a multifunctional protein together with one or more components of the system for biosynthesis of tryptophan (EC 2.4.2.18, EC 4.1.1.48, EC 4.1.3.27, and EC 4.2.1.20)
	Chorismate mutase	
	Tyrosine--tRNA ligase	
	Phenylalanine--tRNA ligase	
Starch and sucrose metabolism	UDP-glucose 6-dehydrogenase	Also acts on UDP-2-deoxyglucose
	Glucoside 3-dehydrogenase	The enzyme acts on D-glucose, D-galactose, D-glucosides and D-galactosides, but D-glucosides react more rapidly than D-galactosides
	CDP-4-dehydro-6-deoxyglucose reductase	Two proteins are involved but no partial reaction has been observed in the presence of

		either alone
	Phosphorylase	The recommended name should be qualified in each instance by adding the name of the natural substance, e.g. maltodextrin phosphorylase, starch phosphorylase, glycogen phosphorylase
	Levansucrase	Some other sugars can act as D-fructosyl acceptors
	Glycogen (starch) synthase	The recommended name varies according to the source of the enzyme and the nature of its synthetic product Glycogen synthase from animal tissues is a complex of a catalytic subunit and the protein glycogenin The enzyme requires glucosylated glycogenin as a primer; this is the reaction product of EC 2.4.1.186 A similar enzyme utilizes ADP-glucose (Cf. EC 2.4.1.21)
	Cellulose synthase (UDP-forming)	Involved in the synthesis of cellulose A similar enzyme utilizes GDP-glucose (Cf. EC 2.4.1.29)
	Sucrose synthase	

	Sucrose-phosphate synthase	
	Alpha,alpha-trehalose-phosphate synthase (UDP-forming)	See also EC 2.4.1.36
	UDP-glucuronosyltransferase	Family of enzymes accepting a wide range of substrates, including phenols, alcohols, amines and fatty acids Some of the activities catalysed were previously listed separately as EC 2.4.1.42, EC 2.4.1.59, EC 2.4.1.61, EC 2.4.1.76, EC 2.4.1.77, EC 2.4.1.84, EC 2.4.1.107 and EC 2.4.1.108 A temporary nomenclature for the various forms whose delineation is in a state of flux
	1,4-alpha-glucan branching enzyme	Converts amylose into amylopectin The recommended name requires a qualification depending on the product, glycogen or amylopectin, e.g. glycogen branching enzyme, amylopectin branching enzyme. The latter has frequently been termed Q-enzyme



	Cellobiose phosphorylase	
	Starch (bacterial glycogen) synthase	The recommended name various according to the source of the enzyme and the nature of its synthetic product, e.g. starch synthase, bacterial glycogen synthase A similar enzyme utilizes UDP-glucose (Cf. EC 2.4.1.11)
	4-alpha-glucanotransferase	An enzymic activity of this nature forms part of the mammalian and Yeast glycogen branching system (see EC 3.2.1.33)
	Cellulose synthase (GDP-forming)	Involved in the synthesis of cellulose A similar enzyme utilizes UDP-glucose (Cf. EC 2.4.1.12)
	1,3-beta-glucan synthase	
	Phenol beta-glucosyltransferase	Acts on a wide range of phenols
	Amylosucrase	
	Polygalacturonate 4-alpha-galacturonosyltransferase	
	Dextranucrase	
	Alpha, alpha-trehalose phosphorylase	
	Sucrose phosphorylase	In the forward reaction, arsenate may replace phosphate In the

		reverse reaction various ketoses and L-arabinose may replace D-fructose
	Maltose phosphorylase	
	1,4-beta-D-xylan synthase	
	Hexokinase	D-glucose, D- mannose, D- fructose, sorbitol and D-glucosamine can act as acceptors ITP and dATP can act as donors The liver isoenzyme has sometimes been called glucokinase
	Phosphoglucoki nase	
	Glucose-1,6- bisphosphate synthase	D-glucose 6- phosphate can act as acceptor, forming D-glucose 1,6- bisphosphate
	Glucokinase	A group of enzymes found in invertebrates and microorganisms highly specific for glucose
	Fructokinase	
	Glucose-1- phosphate phosphodismuta se	
	Protein-N(PI)- phosphohistidine -sugar phosphotransfer ase	Comprises a group of related enzymes The protein substrate is a phosphocarrier protein of low molecular mass (9.5 Kd) A phosphoenzyme intermediate is

		formed The enzyme translocates the sugar it phosphorylates into bacteria Aldohexoses and their glycosides and alditols are phosphorylated on O-6; fructose and sorbose on O-1 Glycerol and disaccharides are also substrates
	Glucose-1-phosphate adenylyltransferase	
	Glucose-1-phosphate cytidylyltransferase	
	Glucose-1-phosphate guanylyltransferase	Also acts, more slowly, on D-mannose 1-phosphate
	UTP--glucose-1-phosphate uridylyltransferase	
	Pectinesterase	
	Trehalose-phosphatase	
	Sucrose-phosphatase	
	Glucose-6-phosphatase	Wide distribution in animal tissues Also catalyses potent transphosphorylations from carbamoyl phosphate, hexose phosphates, pyrophosphate, phosphoenolpyruvate and nucleoside di- and triphosphates,

		to D-glucose, D-mannose, 3-methyl-D-glucose, or 2-deoxy-D-glucose (cf. EC 2.7.1.62, EC 2.7.1.79, and EC 3.9.1.1)
	Alpha-amylase	Acts on starch, glycogen and related polysaccharides and oligosaccharides in a random manner; reducing groups are liberated in the alpha-configuration
	Oligo-1,6-glucosidase	Also hydrolyses palatinose The enzyme from intestinal mucosa is a single polypeptide chain also catalysing the reaction of EC 3.2.1.48
	Maltose-6[PRIME]-phosphate glucosidase	Hydrolyses a variety of 6-phospho-D-glucosides, including maltose 6-phosphate, alpha[PRIME]alpha-trehalose 6-phosphate, sucrose 6-phosphate and p-nitrophenyl-alpha-D-glucopyranoside 6-phosphate (as a chromogenic substrate) The enzyme is activated by Fe(II), Mn(II), Co(II) and Ni(II). It is rapidly inactivated in air
	Polygalacturonase	
	Beta-amylase	Acts on starch,

		glycogen and related polysaccharides and oligosaccharides producing beta-maltose by an inversion
	Alpha-glucosidase	Group of enzymes whose specificity is directed mainly towards the exohydrolysis of 1,4-alpha-glucosidic linkages, and that hydrolyse oligosaccharides rapidly, relative to polysaccharides, which are hydrolysed relatively slowly, or not at all. The intestinal enzyme also hydrolyses polysaccharides, catalysing the reactions of EC 3.2.1.3, and, more slowly, hydrolyses 1,6-alpha-D-glucose links
	Beta-glucosidase	Wide specificity for beta-D-glucosides. Some examples also hydrolyse one or more of the following: beta-D-galactosides, alpha-L-arabinosides, beta-D-xylosides, and beta-D-fucosides
	Beta-fructofuranosidase	Substrates include sucrose. Also catalyses fructotransferase

		reactions
	Alpha,alpha-trehalase	
	Glucan 1,4-alpha-glucosidase	Most forms of the enzyme can rapidly hydrolyse 1,6-alpha-D-glucosidic bonds when the next bond in sequence is 1,4, and some preparations of this enzyme hydrolyse 1,6- and 1,3-alpha-D-glucosidic bonds in other polysaccharides. This entry covers all such enzymes acting on polysaccharides more rapidly than on oligosaccharides EC 3.2.1.20 from mammalian intestine can catalyse similar reactions.
	Beta-glucuronidase	
	Amylo-1,6-glucosidase	In mammals and yeast this enzyme is linked to a glycosyltransferase similar to EC 2.4.1.25; together these two activities constitute the glycogen debranching system.
	Xylan 1,4-beta-xylosidase	Also hydrolyses xylobiose. Some other exoglycosidase activities have been found associated with this enzyme in sheep liver.

	Glucan endo-1,3-beta-D-glucosidase	Very limited action on mixed-link (1,3-1,4-)-beta-D-glucans Hydrolyses laminarin, paramylon and pachyman Different from EC 3.2.1.6
	Cellulase	Will also hydrolyse 1,4-linkages in beta-D-glucans also containing 1,3-linkages
	Sucrose alpha-glucosidase	This enzyme is isolated from intestinal mucosa as a single polypeptide chain also displaying activity towards isomaltose (oligo-1,6-glucosidase, cf. EC 3.2.1.10)
	Cyclomaltodextrinase	Also hydrolyses linear maltodextrin
	Glucan 1,3-beta-glucosidase	Acts on oligosaccharides but very slowly on laminaribiose
	Levanase	
	Galacturan 1,4-alpha-galacturonidase	
	Glucan 1,4-beta-glucosidase	Acts on 1,4-beta-D-glucans and related oligosaccharides Cellobiose is hydrolysed, very slowly
	Cellulose 1,4-beta-cellobiosidase	
	Alpha,alpha-phosphotrehalase	
	ADP-sugar diphosphatase	Has a distinct specificity from the

		UDP-sugar pyrophosphatase (EC 3.6.1.45)
	Nucleotide pyrophosphatase	Substrates include NAD(+), NADP(+), FAD, CoA and also ATP and ADP
	UDP-glucuronate decarboxylase	
	CDP-glucose 4,6-dehydratase	
	CDP-abequose epimerase	
	UDP-glucuronate 4-epimerase	
	Glucose-6-phosphate isomerase	Also catalyses the anomerization of D-glucose 6-phosphate
	Phosphoglucose isomerase	Maximum activity is only obtained in the presence of alpha-D-glucose 1,6-bisphosphate. This bisphosphate is an intermediate in the reaction, being formed by transfer of a phosphate residue from the enzyme to the substrate, but the dissociation of bisphosphate from the enzyme complex is much slower than the overall isomerization. Also, more slowly, catalyses the interconversion of 1-phosphate and 6-phosphate isomers of many other alpha-D-hexoses, and the



		interconversion of alpha-D-ribose 1- phosphate and 5- phosphate
	Beta- phosphoglucom utase	
	Maltose alpha- D- glucosyltransfera se	
Tryptophan metabolism	Indole-3-lactate dehydrogenase	
	Indole-3- acetaldehyde reductase (NADH)	
	Indole-3- acetaldehyde reductase (NADPH)	
	3-hydroxyacyl- CoA dehydrogenase	Also oxidizes S-3- hydroxyacyl-N- acylthioethanolamin e and S-3- hydroxyacylhydrolip oate Some enzymes act, more slowly, with NADP(+) Broad specificity to acyl chain-length (cf. EC 1.1.1.211)
	O-aminophenol oxidase	Isophenoxazine may be formed by a secondary condensation from the initial oxidation product
	Catalase	This enzyme can also act as a peroxidase (EC 1.11.1.7) for which several organic substances, especially ethanol, can act as a

		hydrogen donor A manganese protein containing Mn(III) in the resting state, which also belongs here, is often called pseudocatalase Enzymes from some microorganisms, such as <i>Penicillium simplicissimum</i> , which exhibit both catalase and peroxidase activity, have sometimes been referred to as catalase-peroxidase
	7,8-dihydroxykynurenate 8,8A-dioxygenase	
	Tryptophan 2,3-dioxygenase	Broad specificity towards tryptamine and derivatives including D- and L-tryptophan, 5-hydroxytryptophan and serotonin
	Indole 2,3-dioxygenase	The enzyme from <i>Jasminum</i> is a flavoprotein containing copper, and forms anthranilate as the final product One enzyme from <i>Tecoma stans</i> is also a flavoprotein containing copper and uses three atoms of oxygen per molecule of indole, to form anthranil (3,4-benzisoxazole) A second enzyme from <i>Tecoma stans</i> ,

		which is not a flavoprotein, uses four atoms of oxygen and forms anthranilate as the final product
	2,3-dihydroxyindole 2,3-dioxygenase	
	Indoleamine-pyrrole 2,3-dioxygenase	Acts on many substituted and unsubstituted indoleamines, including melatonin Involved in the degradation of melatonin
	3-hydroxyanthranilate 3,4-dioxygenase	The product of the reaction spontaneously rearrange to quinolinic acid (quin)
	Tryptophan 2-monooxygenase	
	Tryptophan 2[PRIME]-dioxygenase	Acts on a number of indolyl-3-alkane derivatives, oxidizing the 3-side-chain in the 2[PRIME]-position. Best substrates are L-tryptophan and 5-hydroxy-L-tryptophan
	Kynurenine 3-monooxygenase	
	Unspecific monooxygenase	Acts on a wide range of substrates including many xenobiotics, steroids, fatty acids, vitamins and prostaglandins Reactions catalysed include hydroxylation,



	L-amino acid oxidase	
	Amine oxidase (flavin-containing)	Acts on primary amines, and usually also on secondary and tertiary amines
	Amine oxidase (copper-containing)	A group of enzymes including those oxidizing primary amines, diamines and histamine One form of EC 1.3.1.15 from rat kidney also catalyses this reaction
	Acetylindoxyl oxidase	
	Acetylserotonin O-methyltransferase	Some other hydroxyindoles also act as acceptor, more slowly
	Indole-3-pyruvate C-methyltransferase	
	Amine N-methyltransferase	A wide range of primary, secondary, and tertiary amines can act as acceptors, including tryptamine, aniline, nicotine and a variety of drugs and other xenobiotics
	Aralkylamine N-acetyltransferase	Narrow specificity towards aralkylamines, including serotonin Not identical with EC 2.3.1.5
	Acetyl-CoA C-acetyltransferase	
	Tryptophan aminotransferase	Also acts on 5-hydroxytryptophan and, to a lesser

		extent on the phenyl amino acids
	Kynurenine--oxoglutarate aminotransferase	Also acts on 3-hydroxykynurenine
	Thioglucosidase	Has a wide specificity for thioglycosides
	Amidase	
	Formamidase	Also acts, more slowly, on acetamide, propanamide and butanamide
	Arylformamidase	Also acts on other aromatic formylamines
	Nitrilase	Acts on a wide range of aromatic nitriles including (indole-3-yl)-acetonitrile and also on some aliphatic nitriles, and on the corresponding acid amides (cf. EC 4.2.1.84)
	Kynureninase	Also acts on 3[PRIME]-hydroxykynurenine and some other (3-arylcarbonyl)-alanines
	Aromatic-L-amino-acid decarboxylase	Also acts on L-tryptophan, 5-hydroxy-L-tryptophan and dihydroxy-L-phenylalanine (DOPA)
	Phenylpyruvate decarboxylase	Also acts on indole-3-pyruvate
	Aminocarboxymuconate-semialdehyde	The product rearranges non-enzymically to

	decarboxylase	picolinate
	Tryptophanase	Also catalyses the synthesis of tryptophan from indole and serine Also catalyses 2,3-elimination and beta-replacement reactions of some indole-substituted tryptophan analogs of L-cysteine, L-serine and other 3-substituted amino acids
	Enoyl-CoA hydratase	Acts in the reverse direction With cis-compounds, yields (3R)-3-hydroxyacyl-CoA (cf. EC 4.2.1.74)
	Nitrile hydratase	Acts on short-chain aliphatic nitriles, converting them into the corresponding acid amides Does not act on these amides or on aromatic nitriles (cf EC 3.5.5.1)
	Tryptophan--tRNA ligase	
Tyrosine metabolism	Alcohol dehydrogenase	Acts on primary or secondary alcohols or hemiacetals The animal, but not the yeast, enzyme acts also on cyclic secondary alcohols
	(R)-4-hydroxyphenyllactate dehydrogenase	Also acts, more slowly, on (R)-3-phenyllactate, (R)-3-(indole-3-yl)lactate and (R)-lactate
	Hydroxyphenylpyruvate	Also acts on 3-(3,4-dihydroxyphenyl)lact

	reductase	ate Involved with EC 2.3.1.140 in the biosynthesis of rosmarinic acid
	Aryl-alcohol dehydrogenase	A group of enzymes with broad specificity towards primary alcohols with an aromatic or cyclohex-1-ene ring, but with low or no activity towards short-chain aliphatic alcohols
	Catechol oxidase	Also acts on a variety of substituted catechols Many of these enzymes also catalyse the reaction listed under EC 1.14.18.1; this is especially true for the classical tyrosinase
	Iodide peroxidase	
	3,4-dihydroxyphenyl acetate 2,3-dioxygenase	
	4-hydroxyphenylpyruvate dioxygenase	
	Stizolobate synthase	The intermediate product undergoes ring closure and oxidation, with NAD(P)(+) as acceptor, to stizolobic acid
	Stizolobinate synthase	The intermediate product undergoes ring closure and oxidation, with NAD(P)(+) as



		acceptor, to stizolobinic acid
	Gentisate 1,2- dioxygenase	
	Homogentisate 1,2-dioxygenase	
	4- hydroxyphenylac etate 1- monooxygenase	Also acts on 4- hydroxyhydratropate forming 2- methylhomogentisat e and on 4- hydroxyphenoxycet ate forming hydroquinone and glycolate
	4- hydroxyphenylac etate 3- monooxygenase	
	Tyrosine N- monooxygenase	
	Hydroxyphenyla cetonitrile 2- monooxygenase	
	Tyrosine 3- monooxygenase	Activated by phosphorylation, catalysed by EC 2.7.1.128
	Dopamine-beta- monooxygenase	Stimulated by fumarate
	Monophenol monooxygenase	A group of copper proteins that also catalyse the reaction of EC 1.10.3.1, if only 1,2- benzenediols are available as substrate
	Succinate- semialdehyde dehydrogenase (NAD(P)+)	
	Aryl-aldehyde dehydrogenase	Oxidizes a number of aromatic aldehydes, but not aliphatic aldehydes

	Aldehyde dehydrogenase (NAD <sup>+</sup> )	Wide specificity, including oxidation of D-glucuronolactone to D-glucarate
	4-carboxy-2-hydroxymuconate-6-semialdehyde dehydrogenase	Does not act on unsubstituted aliphatic or aromatic aldehydes or glucose NAD(+) can replace NADP(+), but with lower affinity
	Aldehyde dehydrogenase (NAD(P) <sup>+</sup> )	
	4-hydroxyphenylacetaldehyde dehydrogenase	With EC 4.2.1.87, brings about the metabolism of octopamine in <i>Pseudomonas</i>
	Aldehyde oxidase	Also oxidizes quinoline and pyridine derivatives May be identical with EC 1.1.3.22
	L-amino acid oxidase	
	Amine oxidase (flavin-containing)	Acts on primary amines, and usually also on secondary and tertiary amines
	Amine oxidase (copper-containing)	A group of enzymes including those oxidizing primary amines, diamines and histamine One form of EC 1.3.1.15 from rat kidney also catalyses this reaction
	Aralkylamine dehydrogenase	Phenazine methosulfate can act as acceptor Acts on aromatic amines and, more slowly, on

		some long-chain aliphatic amines, but not on methylamine or ethylamine (cf EC 1.4.99.3)
	Phenol O-methyltransferase	Acts on a wide variety of simple alkyl-, methoxy- and halo-phenols
	Tyramine N-methyltransferase	Has some activity on phenylethylamine analogs
	Phenylethanolamine N-methyltransferase	Acts on various phenylethanolamines; converts noradrenalin into adrenalin
	Catechol O-methyltransferase	The mammalian enzymes act more rapidly on catecholamines such as adrenaline or noradrenaline than on catechols
	Glutamine N-phenylacetyltransferase	
	Rosmarinate synthase	Involved with EC 1.1.1.237 in the biosynthesis of rosmarinic acid
	Hydroxymandelonitrile glucosyltransferase	3,4-dihydroxymandelonitrile can also act as acceptor
	Aspartate aminotransferase	Also acts on L-tyrosine, L-phenylalanine and L-tryptophan. This activity can be formed from EC 2.6.1.57 by controlled proteolysis
	Dihydroxyphenylalanine	

	aminotransferase	
	Tyrosine aminotransferase	L-phenylalanine can act instead of L-tyrosine The mitochondrial enzyme may be identical with EC 2.6.1.1 The three isoenzymic forms are interconverted by EC 3.4.22.4
	Aromatic amino acid transferase	L-methionine can also act as donor, more slowly Oxaloacetate can act as acceptor Controlled proteolysis converts the enzyme to EC 2.6.1.1
	Histidinol-phosphate aminotransferase	
	Fumarylacetoacetase	Also acts on other 3,5- and 2,4-dioxo acids
	Acylpyruvate hydrolase	Acts on formylpyruvate, 2,4-dioxopentanoate, 2,4-dioxohexanoate and 2,4-dioxoheptanoate
	Tyrosine decarboxylase	The bacterial enzyme also acts on 3-hydroxytyrosine and, more slowly, on 3-hydroxyphenylalanine
	Aromatic-L-amino-acid decarboxylase	Also acts on L-tryptophan, 5-hydroxy-L-tryptophan and dihydroxy- L-

**Q**uestions **A**nd **A**nswers

	Maleylacetoacetate isomerase	Also acts on maleylpyruvate
	Maleylpyruvate isomerase	
	Phenylpyruvate tautomerase	Also acts on other arylpyruvates
	5-carboxymethyl-2-hydroxymuconate delta-isomerase	
	Tyrosine 2,3-aminomutase	
	Phenylacetate--CoA ligase	Also acts, more slowly, on acetate, propanoate and butanoate, but not on hydroxy derivatives of phenylacetate and related compounds

## IX. PROMOTERS AS SENTINELS

Useful promoters include those that are capable of facilitating preferential transcription, i.e. tissue-specific or developmentally regulated gene expression and being a component of facile systems to evaluate the metabolic/physiological state of a plant cell, tissue or organ. Many such promoters are included in this application. Operably linking a sequence to these promoters that can act as a reporter and inserting the construct into a plant allows detection of the preferential in plantar transcription. For example, the quantitative state of responses to environmental conditions can be detected by using a plant having a construct that contains a stress-inducible promoter linked to and controlling expression of a sequence encoding GFP. The greater the stress promoter is induced, the greater the levels of fluorescence from GFP will be produced and this provides a measure of the level of stress being expressed by the plant and/or the ability of the plant to respond internally to the stress.

More specifically, using this system the activities of any metabolic pathway (catabolic and anabolic), stress-related pathways as on any plant gene repeated activity can be monitored. In addition, assays can be developed using this sentinel system to select for superior genotypes with greater yield characteristics or to select for plants with altered responses to chemical, herbicide, or plant growth regulators or to identify chemical, herbicides or plant growth regulators by their response on such sentinels.

Specifically, a promoter that is regulated in plants in the desired way, is operably linked to a reporter such as GFP, RFP, etc., and the constructs are introduced into the plant of interest. The behavior of the reporter is monitored using technologies typically specific for that reporter. With GFP, RFP, etc., it could typically be by microscopy of whole plants, organs, tissues or cells under excitation by an appropriate wavelength of UV light.

BRIEF DESCRIPTION OF THE TABLES

**BRIEF DESCRIPTION OF THE TABLES**

1. Reference and Sequence Tables

The sequences of exemplary SDFs and polypeptides corresponding to the coding sequences of the instant invention are described in the Reference and Sequence Tables (sometimes referred to as the REF and SEQ Tables. The Reference Table refers to a number of "Maximum Length Sequences" or "MLS." Each MLS corresponds to the longest cDNA obtained, either by cloning or by the prediction from genomic sequence. The sequence of the MLS is the cDNA sequence as described in the Av subsection of the Reference Table.

The Reference Table includes the following information relating to each MLS:

- I. cDNA Sequence
  - A. 5' UTR
  - B. Coding Sequence
  - C. 3' UTR
- II. Genomic Sequence
  - A. Exons
  - B. Introns
  - C. Promoters
- III. Link of cDNA Sequences to Clone IDs
- IV. Multiple Transcription Start Sites
- V. Polypeptide Sequences
  - A. Signal Peptide
  - B. Domains
  - C. Related Polypeptides
- VI. Related Polynucleotide Sequences

I. cDNA SEQUENCE

The Reference Table indicates which sequence in the Sequence Table represents the sequence of each MLS. The MLS sequence can comprise 5' and 3' UTR as well as coding sequences. In addition, specific cDNA clone numbers also are included in the Reference Table when the MLS sequence relates to a specific cDNA clone.

A. 5' UTR

The location of the 5' UTR can be determined by comparing the most 5' MLS sequence with the corresponding genomic sequence as indicated in the Reference Table. The sequence that matches, beginning at any of the transcriptional start sites and ending at the last nucleotide before any of the translational start sites corresponds to the 5' UTR.

B. Coding Region

The coding region is the sequence in any open reading frame found in the MLS. Coding regions of interest are indicated in the PolyP SEQ subsection of the Reference Table.

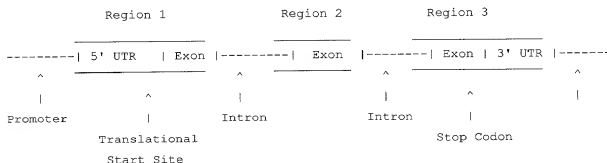
C. 3' UTR



The location of the 3' UTR can be determined by comparing the most 3' MLS sequence with the corresponding genomic sequence as indicated in the Reference Table. The sequence that matches, beginning at the translational stop site and ending at the last nucleotide of the MLS corresponds to the 3' UTR.

## II. GENOMIC SEQUENCE

Further, the Reference Table indicates the specific "gi" number of the genomic sequence if the sequence resides in a public databank. For each genomic sequence, Reference tables indicate which regions are included in the MLS. These regions can include the 5' and 3' UTRs as well as the coding sequence of the MLS. See, for example, the scheme below:



The Reference Table reports the first and last base of each region that are included in an MLS sequence. An example is shown below:

gi No. 47000:

37102 ... 37497

37593 ... 37925

The numbers indicate that the MLS contains the following sequences from two regions of gi No. 47000; a first region including bases 37102-37497, and a second region including bases 37593-37925.

### A. EXON SEQUENCES

The location of the exons can be determined by comparing the sequence of the regions from the genomic sequences with the corresponding MLS sequence as indicated by the Reference Table.

#### i. INITIAL EXON

To determine the location of the initial exon, information from the

- (1) polypeptide sequence section;
- (2) cDNA polynucleotide section; and
- (3) the genomic sequence section

of the Reference Table is used. First, the polypeptide section will indicate where the translational start site is located in the MLS sequence. The MLS sequence can be matched to the genomic sequence that corresponds to the MLS. Based on the match between the MLS and corresponding genomic sequences, the location of the translational start site can be determined in one of the regions of the genomic sequence. The location of this translational start site is the start of the first exon.

Generally, the last base of the exon of the corresponding genomic region, in which the translational start site was located, will represent the end of the initial exon. In some cases, the initial exon will end with a stop codon, when the initial exon is the only exon.

In the case when sequences representing the MLS are in the positive strand of the corresponding genomic sequence, the last base will be a larger number than the first base. When the sequences representing the MLS are in the negative strand of the corresponding genomic sequence, then the last base will be a smaller number than the first base.

#### ii. INTERNAL EXONS

Except for the regions that comprise the 5' and 3' UTRs, initial exon, and terminal exon, the remaining genomic regions that match the MLS sequence are the internal exons. Specifically, the bases defining the boundaries of the remaining regions also define the intron/exon junctions of the internal exons.

#### iii. TERMINAL EXON

As with the initial exon, the location of the terminal exon is determined with information from the

- (1) polypeptide sequence section;
- (2) cDNA polynucleotide section; and
- (3) the genomic sequence section

of the Reference Table. The polypeptide section will indicate where the stop codon is located in the MLS sequence. The MLS sequence can be matched to the corresponding genomic sequence. Based on the match between MLS and corresponding genomic sequences, the location of the stop codon can be determined in one of the regions of the genomic sequence. The location of this stop codon is the end of the terminal exon. Generally, the first base of the exon of the corresponding genomic region that matches the cDNA sequence, in which the stop codon was located, will represent the beginning of the terminal exon. In some cases, the translational start site will represent the start of the terminal exon, which will be the only exon.

In the case when the MLS sequences are in the positive strand of the corresponding genomic sequence, the last base will be a larger number than the first base. When the MLS sequences are in the negative strand of the corresponding genomic sequence, then the last base will be a smaller number than the first base.

#### B. INTRON SEQUENCES

In addition, the introns corresponding to the MLS are defined by identifying the genomic sequence located between the regions where the genomic sequence comprises exons. Thus, introns are defined as starting one base downstream of a genomic region comprising an exon, and end one base upstream from a genomic region comprising an exon.

#### C. PROMOTER SEQUENCES

As indicated below, promoter sequences corresponding to the MLS are defined as sequences upstream of the first exon; more usually, as sequences upstream of the first of multiple transcription start sites; even more usually as sequences about 2,000 nucleotides upstream of the first of multiple transcription start sites.

#### III. LINK of cDNA SEQUENCES to CLONE IDs

As noted above, the Reference Table identifies the cDNA clone(s) that relate to each MLS. The MLS sequence can be longer than the sequences included in the cDNA clones. In such a case, the Reference Table indicates the region of the MLS that is included in the clone. If either the 5' or 3' termini of the cDNA clone sequence is the same as the MLS sequence, no mention will be made.

#### IV. Multiple Transcription Start Sites

Initiation of transcription can occur at a number of sites of the gene. The Reference Table indicates the possible multiple transcription sites for each gene. In the Reference Table, the location of the transcription start sites can be either a positive or negative number.

The positions indicated by positive numbers refer to the transcription start sites as located in the MLS sequence. The negative numbers indicate the transcription start site within the genomic sequence that corresponds to the MLS.

To determine the location of the transcription start sites with the negative numbers, the MLS sequence is aligned with the corresponding genomic sequence. In the instances when a public genomic sequence is referenced, the relevant corresponding genomic sequence can be found by direct reference to the nucleotide sequence indicated by the "gi" number shown in the public genomic DNA section of the Reference Table. When the position is a negative number, the transcription start site is located in the corresponding genomic sequence upstream of the base that matches the beginning of the MLS sequence in the alignment. The negative number is relative to the first base of the MLS sequence which matches the genomic sequence corresponding to the relevant "gi" number.

In the instances when no public genomic DNA is referenced, the relevant nucleotide sequence for alignment is the nucleotide sequence associated with the amino acid sequence designated by "gi" number of the later PolyP SEQ subsection.

#### V. Polypeptide Sequences

The PolyP SEQ subsection lists SEQ ID NOs and Ceres SEQ ID NO for polypeptide sequences corresponding to the coding sequence of the MLS sequence and the location of the translational start site with the coding sequence of the MLS sequence.

The MLS sequence can have multiple translational start sites and can be capable of producing more than one polypeptide sequence.

##### A. Signal Peptide

The Reference tables also indicate in subsection (B) the cleavage site of the putative signal peptide of the polypeptide corresponding to the coding sequence of the MLS sequence. Typically, signal peptide coding sequences comprise a sequence encoding the first residue of the polypeptide to the cleavage site residue.

#### B. Domains

Subsection (C) provides information regarding identified domains (where present) within the polypeptide and (where present) a name for the polypeptide domain.

#### C. Related Polypeptides

Subsection (Dp) provides (where present) information concerning amino acid sequences that are found to be related and have some percentage of sequence identity to the polypeptide sequences of the Reference and Sequence Tables. These related sequences are identified by a "gi" number.

#### VI. Related Polynucleotide Sequences

Subsection (Dn) provides polynucleotide sequences (where present) that are related to and have some percentage of sequence identity to the MLS or corresponding genomic sequence.

Abbreviation	Description
Max Len. Seq.	Maximum Length Sequence
rel to	Related to
Clone Ids	Clone ID numbers
Pub gDNA	Public Genomic DNA
gi No.	gi number
Gen. Seq. in Cdna	Genomic Sequence in cDNA (Each region for a single gene prediction is listed on a separate line. In the case of multiple gene predictions, the group of regions relating to a single prediction are separated by a blank line)
(Ac) cDNA SEQ	cDNA sequence

Abbreviation	Description
- Pat. Appln. SEQ ID NO	Patent Application SEQ ID NO:
- Ceres SEQ ID NO: 1673877	Ceres SEQ ID NO:
- SEQ # w. TSS	Location within the cDNA sequence, SEQ ID NO., of Transcription Start Sites which are listed below
- Clone ID #: # -> #	Clone ID comprises bases # to # of the cDNA Sequence
PolyP SEQ	Polypeptide Sequence
- Pat. Appln. SEQ ID NO:	Patent Application SEQ ID NO:
- Ceres SEQ ID NO	Ceres SEQ ID NO:
- Loc. SEQ ID NO: @ nt.	Location of translational start site in cDNA of SEQ ID NO: at nucleotide number
(C) Pred. PP Nom. & Annot.	Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Title)	Name of Domain
- Loc. SEQ ID NO #: # -> # aa.	Location of the domain within the polypeptide of SEQ ID NO: from # to # amino acid residues.
(Dp) Rel. AA SEQ	Related Amino Acid Sequences
- Align. NO	Alignment number
- gi No	Gi number
- Desp.	Description
- % Idnt.	Percent identity
- Align. Len.	Alignment Length
- Loc. SEQ ID NO: # -> # aa	Location within SEQ ID NO: from # to # amino acid residue.

## 2. Protein Group Table

This table indicates groups of proteins that share a signature sequence (also referred to as a consensus sequence). The Protein group also referred to as the Ortholog group is named by the peptide ID with which all members were compared. Each group contains sequences that were included at the  $10^{-50}$ ,  $10^{-30}$ , and  $10^{-10}$  p-value cutoffs. For each group, the peptide ID and at which cutoff the peptide was included into the group. The same peptide ID may be included in the group three times as peptide ID 50, peptide ID 30 and peptide ID 10. The data indicates that peptide ID was included in the group when the threshold was either  $10^{-50}$ ,  $10^{-30}$ , or  $10^{-10}$ . All the peptide IDs that are followed by "50" were included in the protein group when the e-value cutoff was  $10^{-50}$ . All the peptide IDs that are followed by either "30" or "50" were included in the protein group when the threshold e-value was  $10^{-30}$ . All the peptide IDs that are followed by "10", "30" or "50" were included in the protein group when  $10^{-10}$  was used as the e-value cutoff.

At the end of each protein group is a list of the consensus sequence that proteins share at the  $10^{-50}$ ,  $10^{-30}$ , or  $10^{-10}$ . The consensus sequence contains both lower-case and upper-case letters. The upper-case letters represent the standard one-letter amino acid abbreviations. The lower case letters represent classes of amino acids:

- "t" refers to tiny amino acids, which are specifically alanine, glycine, serine and threonine.
- "p" refers to polar amino acids, which are specifically, asparagine and glutamine
- "n" refers to negatively charged amino acids, which are specifically, aspartic acid and glutamic acid
- "+" refers to positively charged residues, which are specifically, lysine, arginine, and histidine
- "r" refers to aromatic residues, which are specifically, phenylalanine, tyrosine, and tryptophan,
- "a" refers to aliphatic residues, which are specifically, isoleucine, valine, leucine, and methionine

### 3. Protein Group Matrix Table

In addition to each consensus sequence, Applicants have generated a scoring matrix to provide further description of the consensus sequence. The first row of each matrix indicates the residue position in the consensus sequence. The matrix reports number of occurrences of all the amino acids that were found in the group members for every residue position of the signature sequence. The matrix also indicates for each residue position, how many different organisms were found to have a polypeptide in the group that included a residue at the relevant position.

The last line of the matrix indicates all the amino acids that were found at each position of the consensus.

#### **4. MA\_diff Table**

The MA\_diff Table presents the results of the differential expression experiments for the mRNAs, as reported by their corresponding cDNA ID number, that were differentially transcribed under a particular set of conditions as compared to a control sample. The cDNA ID numbers correspond to those utilized in the Reference and Sequence Tables. Increases in mRNA abundance levels in experimental plants versus the controls are denoted with the plus sign (+). Likewise, reductions in mRNA abundance levels in the experimental plants are denoted with the minus (-) sign.

The Table is organized according to each set of experimental conditions, which are denoted by the term "Expt ID:" followed by a particular number. The table below links each Expt ID with a short description of the experiment and the parameters.

For each experiment ID a method of the normalization is specified. "Method: 2" represents normalization by median the goal of the method is to adjust the ratios by a factor so that the median of the ratio distribution is 1. Method 3 is the normalization procedure conducted by Agilent Technologies, Inc. Palo Alto, California, USA.

The MA\_diff Table also specifies the specific parameters and the experiment number (e.g. 107871) used in compiling the data. The experiment numbers are referenced in the appropriate utility/functions sections herein. The background threshold was set to "BKG\_Threshold=X" to reduce the effect of the background on the signal.

Finally, the Table includes reference to an "Organism\_ID" number. This number refers to the cDNA spotted on the chip were similar to *Arabidopsis thaliana* (3769) sequences or whether the oligo used for the chips were similar to *Zea mays* (311987) sequences.

#### **5. MA\_diff (Experiment) Tables**

Example No.	Experiment short name	genome	Expt_ID	Value	Parameter	Units
-------------	--------------------------	--------	---------	-------	-----------	-------



3ii	3642-1	Arabidopsis	108512	3746-1	Plant Line	Hours
3n	Arab_0.001%_MeJA_1	Arabidopsis	108568	Aerial	Tissue	Tissue
				0.001%_MeJ A	Treatment	Compound
				1	Timepoint	Hours
3n	Arab_0.001%_MeJA_1	Arabidopsis	108569	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				0.001%_MeJ A	Treatment	Compound
3j	Arab_0.1uM_E pi-Brass_1	Arabidopsis	108580	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				0.1uM_Brassi no_Steroid	Treatment	Compound
3j	Arab_0.1uM_E pi-Brass_1	Arabidopsis	108581	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				0.1uM_Brassi no_Steroid	Treatment	Compound
3g	Arab_100uM_ ABA_1	Arabidopsis	108560	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				100uM_ABA	Treatment	Compound
3g	Arab_100uM_ ABA_1	Arabidopsis	108561	Aerial	Tissue	Tissue
				100uM_ABA	Treatment	Compound
				6	Timepoint	Hours
3l	Arab_100uM_ BA_1	Arabidopsis	108566	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				100uM_BA	Treatment	Compound
3l	Arab_100uM_ BA_1	Arabidopsis	108567	Aerial	Tissue	Tissue
				100uM_BA	Treatment	Compound
				6	Timepoint	Hours
3k	Arab_100uM_ GA3_1	Arabidopsis	108562	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				100uM GA3	Treatment	Compound
3k	Arab_100uM_ GA3_1	Arabidopsis	108563	Aerial	Tissue	Tissue
				100uM GA3	Treatment	Compound
				6	Timepoint	Hours

3h	Arab_100uM_NAA_1	Arabidopsis	108564	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				100uM_NAA	Treatment	Compound
3h	Arab_100uM_NAA_1	Arabidopsis	108565	Aerial	Tissue	Tissue
				100uM_NAA	Treatment	Compound
				6	Timepoint	Hours
3r	Arab_20%_PE_G_1	Arabidopsis	108570	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				20%PEG	Treatment	Compound
3r	Arab_20%_PE_G_1	Arabidopsis	108571	Aerial	Tissue	Tissue
				20%PEG	Treatment	Compound
				6	Timepoint	Hours
3o	Arab_2mM_SA_1	Arabidopsis	108586	Aerial	Tissue	Tissue
				2mM_SA	Treatment	Compound
				1	Timepoint	Hours
3o	Arab_2mM_SA_1	Arabidopsis	108587	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				2mM_SA	Treatment	Compound
3u	Arab_5mM_H2O2_1	Arabidopsis	108582	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				5mM_H2O2	Treatment	Compound
3u	Arab_5mM_H2O2_1	Arabidopsis	108583	Aerial	Tissue	Tissue
				5mM_H2O2	Treatment	Compound
				6	Timepoint	Hours
3v	Arab_5mM_NaNP_1	Arabidopsis	108584	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				5mM_NaNP	Treatment	Compound
3v	Arab_5mM_NaNP_1	Arabidopsis	108585	Aerial	Tissue	Tissue
				5mM_NaNP	Treatment	Compound
				6	Timepoint	Hours
3t	Arab_Cold_1	Arabidopsis	108578	Aerial	Tissue	Tissue
				Cold	Treatment	Compound
				1	Timepoint	Hours
3t	Arab_Cold_1	Arabidopsis	108579	Aerial	Tissue	Tissue

				6	Timepoint	Hours
				Cold	Treatment	Compound
3g	Arab_Drought_1	Arabidopsis	108572	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				Drought	Treatment	Compound
3g	Arab_Drought_1	Arabidopsis	108573	Aerial	Tissue	Tissue
				Drought	Treatment	Compound
				6	Timepoint	Hours
3s	Arab_Heat_1	Arabidopsis	108576	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				Heat (42 deg C)	Treatment	Compound
3s	Arab_Heat_1	Arabidopsis	108577	Aerial	Tissue	Tissue
				Heat (42 deg C)	Treatment	Compound
				6	Timepoint	Hours
3aa (ovule)	Arab_Ler-pi_ovule_1	Arabidopsis	108595	Ler_pi	Plant Line	Hours
				Ovule	Tissue	Tissue
3b	Arab_Ler-rhl_root_1	Arabidopsis	108594	Ler_rhl	Plant Line	Hours
				Root	Tissue	Tissue
3l	Arab_NO3_H-to-L_1	Arabidopsis	108592	Aerial	Tissue	Tissue
				Low Nitrogen	Treatment	Compound
				12	Timepoint	Hours
3l	Arab_NO3_H-to-L_1	Arabidopsis	108593	Aerial	Tissue	Tissue
				24	Timepoint	Hours
				Low Nitrogen	Treatment	Compound
3l	Arab_NO3_L-to-H_1	Arabidopsis	108588	Aerial	Tissue	Tissue
				2	Timepoint	Hours
				Nitrogen	Treatment	Compound
3l	Arab_NO3_L-to-H_1	Arabidopsis	108589	Aerial	Tissue	Tissue
				Nitrogen	Treatment	Compound
				6	Timepoint	Hours
3l	Arab_NO3_L-to-H_1	Arabidopsis	108590	Aerial	Tissue	Tissue
				9	Timepoint	Hours
				Nitrogen	Treatment	Compound

3l	Arab_NO3_L-to-H_1	Arabidopsis	108591	Aerial	Tissue	Tissue
				Nitrogen	Treatment	Compound
				12	Timepoint	Hours
3p	Arab_Woundin_g_1	Arabidopsis	108574	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				Wounding	Treatment	Compound
3p	Arab_Woundin_g_1	Arabidopsis	108575	Aerial	Tissue	Tissue
				Wounding	Treatment	Compound
				6	Timepoint	Hours
3o	Columbia/CS3726 flower SA	Arabidopsis	108475	Columbia	species	Hours
				SA	Treatment	Compound
				5 weeks	Timepoint	Hours
3o	Columbia/CS3726 flower SA	Arabidopsis	108476	CS3726	species	Hours
				5 weeks	Timepoint	Hours
				SA	Treatment	Compound
3p	Corn_0.001Percent_MeJA	Zea Mays	108555	Aerial	Tissue	Tissue
				24	Timepoint	Hours
				0.001%_MeJA	Treatment	Compound
3j	Corn_0.1uM_Brassinolide	Zea Mays	108557	24	Timepoint	Hours
				Aerial	Tissue	Tissue
				0.1uM_Brassinolide	Treatment	Compound
3g	Corn_100uM_ABA	Zea Mays	108513	Aerial	Tissue	Tissue
				ABA	Treatment	Compound
				6	Timepoint	Hours
3g	Corn_100uM_ABA	Zea Mays	108597	Aerial	Tissue	Tissue
				24	Timepoint	Hours
				100uM_ABA	Treatment	Compound
3i	Corn_100uM_BA	Zea Mays	108517	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				BA	Treatment	Compound
3k	Corn_100uM_GA3	Zea Mays	108519	Aerial	Tissue	Tissue
				100uM	Treatment	Compound



3o	Corn_2mM_SA	Zea Mays	108515	Aerial	Tissue	Tissue
				SA	Treatment	Compound
				12	Timepoint	Hours
3o	Corn_2mM_SA	Zea Mays	108552	Aerial	Tissue	Tissue
				SA	Treatment	Compound
				24	Timepoint	Hours
3u	Corn_5mM_H2O2	Zea Mays	108537	Aerial	Tissue	Tissue
				H2O2	Treatment	Compound
				1	Timepoint	Hours
3u	Corn_5mM_H2O2	Zea Mays	108538	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				H2O2	Treatment	Compound
3u	Corn_5mM_H2O2	Zea Mays	108558	Aerial	Tissue	Tissue
				24	Timepoint	Hours
				H2O2	Treatment	Compound
3v	Corn_5mM_NO	Zea Mays	108526	Aerial	Tissue	Tissue
				NO	Treatment	Compound
				1	Timepoint	Hours
3v	Corn_5mM_NO	Zea Mays	108527	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				NO	Treatment	Compound
3v	Corn_5mM_NO	Zea Mays	108559	Aerial	Tissue	Tissue
				12	Timepoint	Hours
				NO	Treatment	Compound
3t	Corn_Cold	Zea Mays	108533	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				Cold	Treatment	Compound
3t	Corn_Cold	Zea Mays	108534	Aerial	Tissue	Tissue
				Cold	Treatment	Compound
				6	Timepoint	Hours
3q	Corn_Drought	Zea Mays	108502	Drought	Treatment	Compound
				1	Timepoint	Hours
3q	Corn_Drought	Zea Mays	108503	Drought	Treatment	Compound
				6	Timepoint	Hours
3q	Corn_Drought	Zea Mays	108504	Drought	Treatment	Compound
				12	Timepoint	Hours
3q	Corn_Drought	Zea Mays	108556	Drought	Treatment	Compound
				24	Timepoint	Hours

3s	Corn_Heat	Zea Mays	108522	Aerial	Tissue	Tissue
				1	Timepoint	Hours
				Heat (42 deg C)	Treatment	Compound
3s	Corn_Heat	Zea Mays	108523	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				Heat (42 deg C)	Treatment	Compound
3gg	Corn_Imbibed Seeds	Zea Mays	108518	Imbibed	Treatment	Compound
				4	Age	days old
				Roots	Tissue	Tissue
3gg	Corn_Imbibed Seeds	Zea Mays	108528	Imbibed	Treatment	Compound
				Aerial	Tissue	Tissue
				5	Age	days old
3gg	Corn_Imbibed Seeds	Zea Mays	108529	Imbibed	Treatment	Compound
				5	Age	days old
				Root	Tissue	Tissue
3gg	Corn_Imbibed Seeds	Zea Mays	108530	Imbibed	Treatment	Compound
				Aerial	Tissue	Tissue
				6	Age	days old
3gg	Corn_Imbibed Seeds	Zea Mays	108531	Imbibed	Treatment	Compound
				6	Age	days old
				root	Tissue	Tissue
3gg	Corn_Imbibed Seeds	Zea Mays	108545	Imbibed	Treatment	Compound
				Aerial	Tissue	Tissue
				3	Age	days old
3gg	Corn_Imbibed Seeds	Zea Mays	108546	Imbibed	Treatment	Compound
				3	Age	days old
				Root	Tissue	Tissue
3gg	Corn_Imbibed Seeds	Zea Mays	108547	Imbibed	Treatment	Compound
				Aerial	Tissue	Tissue
				4	Age	days old
3gg	Corn_Imbibed_Embryo_Endos perm	Zea Mays	108543	2	Age	days old
				Imbibed	Treatment	Compound
				Embryo	Tissue	Tissue

3gg	Corn_Imbibed_Embryo_Endosperm	Zea Mays	108544	2	Age	days old
				Endosperm	Tissue	Tissue
				Imbibed	Treatment	Compound
3ee	Corn_Meristem	Zea Mays	108535	Root Meristem	Tissue	Tissue
				192	Timepoint	Hours
3ee	Corn_Meristem	Zea Mays	108536	Shoot Meristem	Tissue	Tissue
				192	Timepoint	Hours
3n	Corn_Nitrogen_H_to_L	Zea Mays	108532	Roots	Tissue	Tissue
				Low Nitrogen	Treatment	Compound
				16	Timepoint	Hours
3n	Corn_Nitrogen_H_to_L	Zea Mays	108548	Root	Tissue	Tissue
				Low Nitrogen	Treatment	Compound
				4	Timepoint	Hours
3m	Corn_Nitrogen_L_to_H	Zea Mays	108549	Aerial	Tissue	Tissue
				0.166	Timepoint	Hours
				Nitrogen	Treatment	Compound
3m	Corn_Nitrogen_L_to_H	Zea Mays	108550	Aerial	Tissue	Tissue
				Nitrogen	Treatment	Compound
				1.5	Timepoint	Hours
3m	Corn_Nitrogen_L_to_H	Zea Mays	108551	Aerial	Tissue	Tissue
				3	Timepoint	Hours
				Nitrogen	Treatment	Compound
3ff	Corn_RT1	Zea Mays	108599	Unknown	Plant Line	Hours
				Root	Tissue	Tissue
3p	Corn_Woundin_g	Zea Mays	108524	Aerial	Tissue	Tissue
				Wounding	Treatment	Compound
				1	Timepoint	Hours
3p	Corn_Woundin_g	Zea Mays	108525	Aerial	Tissue	Tissue
				6	Timepoint	Hours
				Wounding	Treatment	Compound
3g	Drought_Flowers	Arabidopsis	108473	Flowers	Tissue	Tissue
				7 d	Timepoint	Hours
				Drought	Treatment	Compound



3g	Drought_Flowers	Arabidopsis	108474	Flowers	Tissue	Tissue
				Drought	Treatment	Compound
				8 d (1d-post_re-watering)	Timepoint	Hours
3k	GA Treated	Arabidopsis	108484	1	Timepoint	Hours
				1	Timepoint	Hours
3k	GA Treated	Arabidopsis	108485	6	Timepoint	Hours
				6	Timepoint	Hours
3k	GA Treated	Arabidopsis	108486	12	Timepoint	Hours
				12	Timepoint	Hours
3e	Germinating Seeds	Arabidopsis	108461	Day 1	Timepoint	Hours
3e	Germinating Seeds	Arabidopsis	108462	Day 2	Timepoint	Hours
3e	Germinating Seeds	Arabidopsis	108463	Day 3	Timepoint	Hours
3e	Germinating Seeds	Arabidopsis	108464	Day 4	Timepoint	Hours
3bb	Herbicide V3.1	Arabidopsis	108465	Round up	Treatment	Compound
				12	Timepoint	Hours
3bb	Herbicide V3.1	Arabidopsis	108466	Trimec	Treatment	Compound
				12	Timepoint	Hours
3bb	Herbicide V3.1	Arabidopsis	108467	Finale	Treatment	Compound
				12	Timepoint	Hours
3bb	Herbicide V3.1	Arabidopsis	108468	Glean	Treatment	Compound
				12	Timepoint	Hours
3bb	Herbicide_v2	Arabidopsis	107871	Finale	Treatment	Compound
				4	Timepoint	Hours
3bb	Herbicide_v2	Arabidopsis	107876	Finale	Treatment	Compound
				12	Timepoint	Hours
3bb	Herbicide_v2	Arabidopsis	107881	Glean	Treatment	Compound
				4	Timepoint	Hours
3bb	Herbicide_v2	Arabidopsis	107886	Trimec	Treatment	Compound
				4	Timepoint	Hours
3bb	Herbicide_v2	Arabidopsis	107891	Trimec	Treatment	Compound
				12	Timepoint	Hours
3bb	Herbicide_v2	Arabidopsis	107896	Round-up	Treatment	Compound
				4	Timepoint	Hours
3d	Trichome Inflorescences expt	Arabidopsis	108452	Hairy Inflorescence #1	Tissue	Tissue
3o	SA treatment_1	Arabidopsis	108471	Columbia	Species	Hours

	hour					
				1	Timepoint	Hours
				SA	Treatment	Compound
3o	SA treatment_1 hour	Arabidopsis	108472	CS3726	Species	Hours
				1	Timepoint	Hours
				SA	Treatment	Compound
3o	SA treatment_4 hour	Arabidopsis	108469	columbia	Species	Hours
				4	Timepoint	Hours
				SA	Treatment	Compound
3o	SA treatment_4 hour	Arabidopsis	108470	CS3726	Species	Hours
				SA	Treatment	Compound
				4	Timepoint	Hours
3o	SA treatment_AJ	Arabidopsis	107953	50	Probe Amount	% of Standard Amount
				SA	Treatment	Compound
				24	Timepoint	Hours
				Clontech	Probe Type	Probe method
3o	SA treatment_AJ	Arabidopsis	107960	50	Probe Amount	% of Standard Amount
				SA	Treatment	Compound
				24	Timepoint	Hours
				Operon	Probe Type	Probe method
3o	SA_treatment 24 hour	Arabidopsis	108443	SA	Treatment	Compound
				24	Timepoint	Hours
3o	SA_treatment 6 hour	Arabidopsis	108440	SA treatment 6 hour	Treatment	Compound
				CS3726	species	Hours
3o	SA_treatment 6 hour	Arabidopsis	108441	SA treatment 6 hour	Treatment	Compound
				Columbia	species	Hours
3l	Nitrogen High transition to Low	Arabidopsis	108454	10 min	Timepoint	Hours
3l	Nitrogen High transition to Low	Arabidopsis	108455	1 hr	Timepoint	Hours
3j	BR Shoot	Arabidopsis	108478	dwf4-1	Plant Line	Hours

	Apices Expt					
3j	BR Shoot Apices Expt	Arabidopsis	108479	AOD4-4	Plant Line	Hours
3j	BR Shoot Apices Expt	Arabidopsis	108480	Ws-2	Plant Line	Hours
				BL	Treatment	Compound
3j	BR Shoot Apices Expt	Arabidopsis	108481	Ws-2	Plant Line	Hours
				BRZ	Treatment	Compound
3jj	Tissue Specific Expression	Arabidopsis	108429	green flower	Tissue	Tissue
				operon	Probe Type	Probe method
				50	Probe Amount	% of Standard Amount
3jj	Tissue Specific Expression	Arabidopsis	108430	white flower	Tissue	Tissue
				50	Probe Amount	% of Standard Amount
				operon	Probe Type	Probe method
3jj	Tissue Specific Expression	Arabidopsis	108431	flowers (bud)	Tissue	Tissue
				operon	Probe Type	Probe method
				50	Probe Amount	% of Standard Amount
3c	Tissue Specific Expression	Arabidopsis	108436	5-10mm siliques	Tissue	Tissue
				33	Probe Amount	% of Standard Amount
				operon	Probe Type	Probe method
3c	Tissue Specific Expression	Arabidopsis	108437	<5mm siliques	Tissue	Tissue
				operon	Probe Type	Probe method
				33	Probe Amount	% of Standard Amount
3c	Tissue Specific	Arabidopsis	108438	5wk siliques	Tissue	Tissue

	Expression					
				33	Probe Amount	% of Standard Amount
				operon	Probe Type	Probe method
3a	Tissue Specific Expression	Arabidopsis	108439	Roots (2wk)	Tissue	Tissue
				operon	Probe Type	Probe method
				33	Probe Amount	% of Standard Amount
3c	Tissue Specific Expression	Arabidopsis	108497	3 week Rossette leaves	Tissue	Tissue
				100	Probe Amount	% of Standard Amount
				operon	Probe Type	Probe method
3c	Tissue Specific Expression	Arabidopsis	108498	3-week stems	Tissue	Tissue
				operon	Probe Type	Probe method
				100	Probe Amount	% of Standard Amount
3dd	U.A.E. Knockout	Arabidopsis	108451	13B12	Plant Line	Hours
3q	Ws Arabidopsis Drought 2 days	Arabidopsis	108477	stems and leaves	Tissue	Tissue
				2 days	Timepoint	Hours
3q	Ws Arabidopsis Drought 4 days	Arabidopsis	108482	4 days	Timepoint	Hours
3q	Ws Arabidopsis Drought 6 days	Arabidopsis	108483	6 days	Timepoint	Hours
3cc	ap2-floral buds	Arabidopsis	108501	ap2 (Ler.)	Plant Line	Hours
				floral buds	Tissue	Tissue
3m	nitrogen-seed set	Arabidopsis	108487	0.5	Timepoint	Hours
3m	nitrogen-seed set	Arabidopsis	108488	2	Timepoint	Hours
3m	nitrogen-seed set	Arabidopsis	108489	4	Timepoint	Hours

3b	rhl mutant2	Arabidopsis	108433	mutant	Tissue	Tissue
3ee	root tips	Arabidopsis	108434	root tips	Tissue	Tissue
3f	stm mutants	Arabidopsis	108435	stem	Tissue	Tissue
	Aluminum		SMD 7304, SMD 7305			
	Axel		SMD 6654, SMD 6655			
	Cadium		SMD 7427, SMD 7428			
	Cauliflower		SMD 5329, SMD 5330			
	Chloroplast		SMD 8093, SMD 8094			
	Circadian		SMD 2344, SMD 2359, SMD 2361, SMD 2362, SMD 2363, SMD 2364, SMD 2365, SMD 2366, SMD 2367, SMD 2368, SMD 3242			
	CO2		SMD7561, SMD 7562, SMD 7261, SMD			

			7263, SMD 3710, SMD 4649, SMD 4650			
	Disease		SMD 7342, SMD 7343			
	reactive oxygen		SMD 7523			
	Iron		SMD 7114, SMD 7115, SMD 7125			
	defense		SMD 8031, SMD 8032			
	Mitochondria- Electron Transport		SMD 8061, SMD 8063			
	NAA		SMD 3743, SMD 3749, SMD 6338, SMD 6339			
	Nitrogen		SMD 3787, SMD 3789			
	Phototropism		SMD 4188, SMD 6617, SMD 6619			
	Shade		SMD 8130, SMD 7230			
	Sqn		SMD 7133, SMD 7137			
	Sulfur		SMD 8034, SMD 8035			

	Wounding		SMD 3714, SMD 3715			
	Zinc		SMD 7310, SMD 7311			

## 6. MA\_Clusters Table

Microarray data was clustered using one of two methods: “complete linkage” or “nearest neighbor” analysis. These clustering methods are described in more detail elsewhere herein. The results of the clustering analysis are presented in the MA\_clust table. The table is organized as follows:

“METHOD” refers to a method number which clustering method used.

“CL\_METHOD\_TYPE=TRUE” refers to complete linkage method.

“NN\_METHOD\_TYPE=TRUE” refers to the nearest neighbor method.

“FULL\_NN\_METHOD\_TYPE=TRUE” refers to the nearest neighbor method, where no size limitation was placed on the cluster.

“PARAMETERS” refers to the parameters utilized for the analysis. The nature of these is also described in more detail elsewhere herein.

“ORGANISM” refers to the cDNA spotted on the chip were similar to Arabidopsis thaliana (3769) sequences or whether the oligo used for the chips were similar to Zea mays (311987) sequences.

Each cluster or group of cDNA is identified by a “Group #”, following which are the individual cDNA\_Ids that are a member of that Group

## 7. Knock-in Table

The Knock-In Table presents the results of knock-in experiments wherein plants are grown from tissues transformed with a marker gene-containing insert and phenotypes are ascertained from the transformed plants. Each section of the Table relating to information on a new transformant begins with a heading “Knock-in phenotype in gene (cDNA\_id):” followed by a number which represents the Ceres internal code for a proprietary cDNA sequence. The described transformant was prepared by procedures described herein, wherein the identified

Ceres proprietary cDNA\_id (corresponding to the cDNA\_id in the Reference and Sequence Tables) was interrupted by the marker gene-containing insert. The following information is presented for each section.

- Parent plants used in cross – presents the id numbers of the parent plants which were crossed to produce the F1 generation plant for which a phenotype is described. The parent plant with the promoter is described by a plant line descriptor.
- Clone ID – presents the clone number of the Ceres proprietary clone which was the source of the cDNA\_id.
- Phenotype ID – represents an internal identification code.
- Unique F1 plant ID – represents the internal code for the F1 plant for which a phenotype is described.
- Assay – presents the type of growth analyzed (e.g. soil gross morphology), followed by the assay name which corresponds to the type/location of the tissue that was observed., the name of the assay conducted for which the result provided the identified phenotype.
- Phenotype – describes the phenotype noted for the F1 generation transformant.
- Notes – may provide additional information on the described phenotype for the transformant.

Each knock-in representing a transformant with an interruption in the identified cDNA\_id may be correlated with more than one identified phenotype.

## **8. Knock-out Table**

The Knock-Out Table presents the results of knock-out experiments wherein plants are grown from tissues transformed with a marker gene-containing insert wherein phenotypes are ascertained from the transformed plants. Each section of the Table relating to information on a new transformant begins with a heading “tail id:” representing an internal code. The following information is presented for each section.

*br* - provides another internal code for the experiment.



*Phenotype\_id* - provides an identification number for the particular phenotype identified for the transformant.

*assay* - identifies the assay procedure utilized in the experiment to identify a phenotype for the transformant.

*phenotype* - represents an internal identification code.

*ratio* - represents a segregation ratio.

*notes* - lists any notes relevant to the identified phenotype.

*Knock-out in-genes* - Identifies the genes in which the tag has inserted

- 6) the less than 501 upstream of the transcriptional start site;
- 7) less than 701 upstream of the translational initiation codon;
- 8) between the translational initiation and termination codons of the gene,
- 9) less than 301 downstream of the translational stop codon; or
- 10) less than 151 downstream of a transcriptional termination site or a gene.

In this table the gene is identified by its cDNA ID number, the Ceres SEQ ID that is indicated in the (Ac) portion of the Reference tables. For each cDNA\_id, the following information is provided:

- the cDNA\_id number.
- in parenthesis, the cluster number of which the identified cDNA is a member.
- the "gDNA\_Insert pos" representing the position of the insert in the corresponding gDNA sequence
- the gi number refers to the TIGR chromosome sequences for Arabidopsis.

*Knock-out out-of-genes:* Identifies the Ceres cDNA proprietary sequences (noted by cDNA\_id which are the same as those identified in the Reference and Sequence Tables) which are closest in position to the insert, both upstream and downstream from the insert. For each cDNA\_id, the following information is provided:

- In the first parentheses, R indicates that the gene is to the right of the tag, L indicates that the gene is to right of the tag as the sequences is read left to right
- the cDNA\_id number

- in next parentheses, the cluster number of which the identified cDNA is a member.
- the distance (in number of nucleotides) of the insert is upstream of the start of the gene annotation as described in the Reference Tables or downstream at the end the gene annotation.
- the “gDNA\_Insert pos” representing the position of the insert in the corresponding gDNA sequence
- the gi number refers to the TIGR chromosome sequences for Arabidopsis.

## 9. Protein Domain Table

The Protein Domain table provides details concerning the protein domains noted in the Reference Table. The majority of the protein domain descriptions given in the Protein Domain Table are obtained from Prosite, (<http://www.expasy.ch/prosite/>), and Pfam, (<http://pfam.wustl.edu/browse.shtml>). Each description in The Table begins with the pfam and Prosite identifying numbers, the full name of the domain, and a detailed description, including biological and in vivo implications/functions for the domain, references which further describe such implications/functions, and references that describe tests/assays to measure the implications/functions.

## 10. Single Gene Functions & Utilities Table

The Single Gene Functions & Utilities Table describes particular utilities/functions of interest for individual genes. The Table identifies the cDNA\_ID of interest, correlates to that cDNA the relevant phenotype, protein domain and microarray/differential expression data. The final column of the Table identifies the utilities/functions of particular interest for the identified cDNA.

## 11. Cluster Functions & Utilities Table

The Cluster Functions & Utilities Table describes particular utilities/functions of interest for identified clusters of genes. The Table provides the following information:

Record # - an internal identifier.

Group – identifies the group of clusters of interest, wherein each group is identified with the same utilities/functions as set forth in the right-hand most column.

CDNA – identifies the cDNA of interest with the noted utility/function.

CDNA\_Cluster - identifies the cDNA Cluster ID of interest.

Gi No – refers to the public genomic sequence that matches to the cDNA

NR Hit – refers to the most relevant protein domain for the cDNA of interest.

Pfam and Pfam Desc – provide the protein domain name.

Notes/Annotations – provides some notes relevant to the data/information analysis.

Utilities/Functions – this rightmost column identifies utilities/functions of particular interest for the group of cDNAs and clusters.

#### 12. cDNA Clusters Table

The cDNA\_Clusters Table correlates the Ceres cDNA\_ID nos. (in numerical order) with the relevant cDNA cluster which contains each cDNA\_ID.

#### 13. Stanford\_old\_new\_cDNA\_map Table

During the course of the experiments reported herein, some of the cDNA sequences were assigned new Ceres internal cDNA\_id numbers. The cDNA\_map Table provides a list of the original “old” cDNA\_ids and correlates those id numbers with any new cDNA\_id which may have been assigned. Thus, any “old” and “new” cDNA ids which are on the same line in the Table are, in fact, the same sequence.

#### 14. gb\_Only\_Peptides Table

In the Protein Group table, a number of proteins encoded by Genbank predictions are included. These proteins were referenced with a peptide ID number. The peptide ID number is linked to the amino acid sequence of the Genbank prediction in this table.

#### 15. Stanford\_Old\_New\_cDNA Table

During the course of the experiments reported herein, some of the cDNA sequences utilized in the Stanford Microarray differential expression analysis experiments were assigned

new Ceres internal cDNA\_id numbers. The Stanford\_old\_new\_cDNA Table provides a list of the original "old" cDNA\_ids and correlates those id numbers with any new cDNA\_id which may have been assigned. Thus, any "old" and "new" cDNA ids which are on the same line in the Table are, in fact, the same sequence.

16. Enhanced Amino Table

This table list the peptide IDs of polypeptides with enhanced amino acid content. The table list the peptide ID following with the single letter code of the amino acid that is enhanced. The table also includes a frequency that the amino acid occurred. The frequency was calculated by dividing the total number of the desired amino acid indicated in the column by the number of residues in the peptide. For example, if amino acid A, occurred 50 times in a polypeptide that is 100 amino acid long, the frequency would be 50 divided by 100 or 0.5.

17. Stanford\_old\_new\_cDNA\_map Table\*\*Keep only for patent #3

During the course of the experiments reported herein, some of the cDNA sequences were assigned new Ceres internal cDNA\_id numbers. The docket\_80090\_101\_cDNA\_map provides a list of the original "old" cDNA\_ids in the Reference and Sequence tables and correlates those id numbers with any new cDNA\_id which may have been assigned and utilized in the remaining tables. Thus, any "old" and "new" cDNA ids which are on the same line in the Table are, in fact, the same sequence.

17. The AFLP\_int Table

The table includes the following headings:

- A. Band\_id
- B. Mobility
- C. S1-S11

A. Band\_id

This text is the band identified in the experiment.

- C. B. Mobility
- D. The mobility of the nucleic acid bands in relation to markers in the gel.

C. S1-S11

These are the various tissue samples from *Arabidopsis thaliana* (wassilewskija) listed below:

- S1 Dark adapted seedlings
- S2 Roots/etiolated seedlings
- S3 Mature leafs, soil grown
- S4 Immature buds, inflorescence meristem
- S5 Flowers opened
- S6 Siliques, all stages
- S7 Senescing leaves (just beginning to yellow)
- S8 A pooled sample of:
  - Callus Inducing medium
  - Callus shoot induction
  - Callus root induction
- S9 Wounding
  - Methyl-jasmonate-treated
- S10 Oxidative stress
  - Drought stress
  - Oxygen Stress-flooding
- S11 Heat treated light grown seedling
  - Cold treated light grown seedlings

The numbers below each sample correspond to the levels of expression of a particular gene, polynucleotide, or nucleic acid. The stronger the intensity of the band on a AFLP gel, the higher the levels of expression.

19. The AFLP\_diff Table

This table shows which band\_id corresponds to which specific cDNA Ceres SEQ ID

NOs. Differential expression of the band is shown by a "+" or "-".



HOW TO MAKE DIFFERENT EMBODIMENTS OF THE INVENTION

The invention relates to (I) polynucleotides and methods of use thereof, such as

- IA. Probes, Primers and Substrates;
- IB. Methods of Detection and Isolation;
  - B.1. Hybridization;
  - B.2. Methods of Mapping;
  - B.3. Southern Blotting;
  - B.4. Isolating cDNA from Related Organisms;
  - B.5. Isolating and/or Identifying Orthologous Genes
- IC. Methods of Inhibiting Gene Expression
  - C.1. Antisense
  - C.2. Ribozyme Constructs;
  - C.3. Chimeraplasts;
  - C.4. Co-Suppression;
  - C.5. Transcriptional Silencing
  - C.6. Other Methods to Inhibit Gene Expression
- ID. Methods of Functional Analysis;
- IE. Promoter Sequences and Their Use;
- IF. UTRs and/or Intron Sequences and Their Use; and
- IG. Coding Sequences and Their Use.

The invention also relates to (II) polypeptides and proteins and methods of use thereof, such

as

- IIA. Native Polypeptides and Proteins
  - A.1 Antibodies
  - A.2 In Vitro Applications
- IIIB. Polypeptide Variants, Fragments and Fusions
  - B.1 Variants
  - B.2 Fragments
  - B.3 Fusions

The invention also includes (III) methods of modulating polypeptide production, such as

IIIA. Suppression

A.1 Antisense

A.2 Ribozymes

A.3 Co-suppression

A.4 Insertion of Sequences into the Gene to be Modulated

A.5 Promoter Modulation

A.6 Expression of Genes containing Dominant-Negative Mutations

IIIB. Enhanced Expression

B.1 Insertion of an Exogenous Gene

B.2 Promoter Modulation

The invention further concerns (IV) gene constructs and vector construction, such as

IVA. Coding Sequences

IVB. Promoters

IVC. Signal Peptides

The invention still further relates to

V. Transformation Techniques

I. Polynucleotides

Exemplified SDFs of the invention represent fragments of the genome of corn, wheat, rice, soybean or *Arabidopsis* and/or represent mRNA expressed from that genome. The isolated nucleic acid of the invention also encompasses corresponding fragments of the genome and/or cDNA complement of other organisms as described in detail below.

Polynucleotides of the invention can be isolated from polynucleotide libraries using primers comprising sequence similar to those described by or Reference, Sequence or polynucleotides that encode sequences Protein Group, and Protein Group Matrix tables or complements thereof. See, for example, the methods described in Sambrook *et al.*, *supra*.

Alternatively, the polynucleotides of the invention can be produced by chemical synthesis. Such synthesis methods are described below.



It is contemplated that the nucleotide sequences presented herein may contain some small percentage of errors. These errors may arise in the normal course of determination of nucleotide sequences. Sequence errors can be corrected by obtaining seeds deposited under the accession numbers cited above, propagating them, isolating genomic DNA or appropriate mRNA from the resulting plants or seeds thereof, amplifying the relevant fragment of the genomic DNA or mRNA using primers having a sequence that flanks the erroneous sequence, and sequencing the amplification product.

#### I.A. Probes, Primers and Substrates

SDFs of the invention can be applied to substrates for use in array applications such as, but not limited to, assays of global gene expression, for example under varying conditions of development, growth conditions. The arrays can also be used in diagnostic or forensic methods (WO95/35505, US 5,445,943 and US 5,410,270).

Probes and primers of the instant invention will hybridize to a polynucleotide comprising a sequence in or encoded by those in the Reference, Sequence, Protein Group, and Protein Group Matrix tables or fragments or complement thereof. Though many different nucleotide sequences can encode an amino acid sequence, the sequences of the reference and Sequence table or sequences that encode polypeptides or fragments thereof described in Protein Group and Protein Group Matrix tables are generally preferred for encoding polypeptides of the invention.

However, the sequence of the probes and/or primers of the instant invention need not be identical to those in the Reference and Sequence tables or the complements thereof. For example, some variation in probe or primer sequence and/or length can allow additional family members to be detected, as well as orthologous genes and more taxonomically distant related sequences. Similarly, probes and/or primers of the invention can include additional nucleotides that serve as a label for detecting the formed duplex or for subsequent cloning purposes.

Probe length will vary depending on the application. For use as primers, probes are 12-40 nucleotides, preferably 18-30 nucleotides long. For use in mapping, probes are preferably 50 to 500 nucleotides, preferably 100-250 nucleotides long. For Southern hybridizations, probes as long as several kilobases can be used as explained below.

The probes and/or primers can be produced by synthetic procedures such as the triester method of Matteucci et al. *J. Am. Chem. Soc.* 103:3185(1981); or according to Urdea et al. *Proc.*

Natl. Acad. 80:7461 (1981) or using commercially available automated oligonucleotide synthesizers.

#### I.B. Methods of Detection and Isolation

The polynucleotides of the invention can be utilized in a number of methods known to those skilled in the art as probes and/or primers to isolate and detect polynucleotides, including, without limitation: Southern, Northern, Branched DNA hybridization assays, polymerase chain reaction, and microarray assays, and variations thereof. Specific methods given by way of examples, and discussed below include:

Hybridization

Methods of Mapping

Southern Blotting

Isolating cDNA from Related Organisms

Isolating and/or Identifying Orthologous Genes.

Also, the nucleic acid molecules of the invention can be used in other methods, such as high density oligonucleotide hybridizing assays, described, for example, in U.S. Pat. Nos. 6,004,753; 5,945,306; 5,945,287; 5,945,308; 5,919,686; 5,919,661; 5,919,627; 5,874,248; 5,871,973; 5,871,971; and 5,871,930; and PCT Pub. Nos. WO 9946380; WO 9933981; WO 9933870; WO 9931252; WO 9915658; WO 9906572; WO 9858052; WO 9958672; and WO 9810858.

#### B.1. Hybridization

The isolated SDFs of the Reference and Sequence tables or SDFs encoding polypeptides of the Protein Group and Protein Group Matrix tables or fragments thereof of the present invention can be used as probes and/or primers for detection and/or isolation of related polynucleotide sequences through hybridization. Hybridization of one nucleic acid to another constitutes a physical property that defines the subject SDF of the invention and the identified related sequences. Also, such hybridization imposes structural limitations on the pair. A good general discussion of the factors for determining hybridization conditions is provided by Sambrook et al. ("Molecular Cloning, a Laboratory Manual, 2nd ed., c. 1989 by Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; *see esp.*, chapters 11 and 12). Additional

considerations and details of the physical chemistry of hybridization are provided by G.H. Keller and M.M. Manak "DNA Probes", 2<sup>nd</sup> Ed. pp. 1-25, c. 1993 by Stockton Press, New York, NY.

Depending on the stringency of the conditions under which these probes and/or primers are used, polynucleotides exhibiting a wide range of similarity to those in the Reference and Sequence or encoding polypeptides of the Protein Group and Protein Group Matrix tables or fragments thereof can be detected or isolated. When the practitioner wishes to examine the result of membrane hybridizations under a variety of stringencies, an efficient way to do so is to perform the hybridization under a low stringency condition, then to wash the hybridization membrane under increasingly stringent conditions.

When using SDFs to identify orthologous genes in other species, the practitioner will preferably adjust the amount of target DNA of each species so that, as nearly as is practical, the same number of genome equivalents are present for each species examined. This prevents faint signals from species having large genomes, and thus small numbers of genome equivalents per mass of DNA, from erroneously being interpreted as absence of the corresponding gene in the genome.

The probes and/or primers of the instant invention can also be used to detect or isolate nucleotides that are "identical" to the probes or primers. Two nucleic acid sequences or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence as described below.

Isolated polynucleotides within the scope of the invention also include allelic variants of the specific sequences presented in the Reference, Sequence, Protein Group, and Protein Group Matrix tables. The probes and/or primers of the invention can also be used to detect and/or isolate polynucleotides exhibiting at least 80% sequence identity with the sequences of the reference, Sequence or encoding polypeptides of the Protein Group and Protein Group Matrix tables or fragments thereof.

With respect to nucleotide sequences, degeneracy of the genetic code provides the possibility to substitute at least one base of the base sequence of a gene with a different base without causing the amino acid sequence of the polypeptide produced from the gene to be changed. Hence, the DNA of the present invention may also have any base sequence that has been changed from a sequence in the Reference, Sequence, Protein Group, and Protein Group

Matrix tables by substitution in accordance with degeneracy of genetic code. References describing codon usage include: Carels *et al.*, *J. Mol. Evol.* 46: 45 (1998) and Fennoy *et al.*, *Nucl. Acids Res.* 21(23): 5294 (1993).

5           B.2.   Mapping

10           The isolated SDF DNA of the invention can be used to create various types of genetic  
15           and physical maps of the genome of corn, Arabidopsis, soybean, rice, wheat, or other plants.  
20           Some SDFs may be absolutely associated with particular phenotypic traits, allowing construction  
25           of gross genetic maps. While not all SDFs will immediately be associated with a phenotype, all  
30           SDFs can be used as probes for identifying polymorphisms associated with phenotypes of  
35           interest. Briefly, one method of mapping involves total DNA isolation from individuals. It is  
40           subsequently cleaved with one or more restriction enzymes, separated according to mass,  
45           transferred to a solid support, hybridized with SDF DNA and the pattern of fragments compared.  
50           Polymorphisms associated with a particular SDF are visualized as differences in the size of  
55           fragments produced between individual DNA samples after digestion with a particular restriction  
60           enzyme and hybridization with the SDF. After identification of polymorphic SDF sequences,  
65           linkage studies can be conducted. By using the individuals showing polymorphisms as parents in  
70           crossing programs, F2 progeny recombinants or recombinant inbreds, for example, are then  
75           analyzed. The order of DNA polymorphisms along the chromosomes can be determined based  
80           on the frequency with which they are inherited together versus independently. The closer two  
85           polymorphisms are together in a chromosome the higher the probability that they are inherited  
90           together. Integration of the relative positions of all the polymorphisms and associated marker  
95           SDFs can produce a genetic map of the species, where the distances between markers reflect the  
100          recombination frequencies in that chromosome segment.

105          The use of recombinant inbred lines for such genetic mapping is described for  
110          *Arabidopsis* by Alonso-Blanco *et al.* (*Methods in Molecular Biology*, vol.82, "Arabidopsis  
115          Protocols", pp. 137-146, J.M. Martinez-Zapater and J. Salinas, eds., c. 1998 by Humana Press,  
120          Totowa, NJ) and for corn by Burr ("Mapping Genes with Recombinant Inbreds", pp. 249-254.  
125          *In Freeling, M. and V. Walbot (Ed.), The Maize Handbook*, c. 1994 by Springer-Verlag New  
130          York, Inc.: New York, NY, USA; Berlin Germany; Burr *et al. Genetics* (1998) 118: 519;

Gardiner, J. et al., (1993) *Genetics* 134: 917). This procedure, however, is not limited to plants and can be used for other organisms (such as yeast) or for individual cells.

The SDFs of the present invention can also be used for simple sequence repeat (SSR) mapping. Rice SSR mapping is described by Morgante et al. (*The Plant Journal* (1993) 3: 165), Panaud et al. (*Genome* (1995) 38: 1170); Senior et al. (*Crop Science* (1996) 36: 1676), Taramino et al. (*Genome* (1996) 39: 277) and Ahn et al. (*Molecular and General Genetics* (1993) 241: 483-90). SSR mapping can be achieved using various methods. In one instance, polymorphisms are identified when sequence specific probes contained within an SDF flanking an SSR are made and used in polymerase chain reaction (PCR) assays with template DNA from two or more individuals of interest. Here, a change in the number of tandem repeats between the SSR-flanking sequences produces differently sized fragments (U.S. Patent 5,766,847). Alternatively, polymorphisms can be identified by using the PCR fragment produced from the SSR-flanking sequence specific primer reaction as a probe against Southern blots representing different individuals (U.H. Refseth et al., (1997) *Electrophoresis* 18: 1519).

Genetic and physical maps of crop species have many uses. For example, these maps can be used to devise positional cloning strategies for isolating novel genes from the mapped crop species. In addition, because the genomes of closely related species are largely syntenic (that is, they display the same ordering of genes within the genome), these maps can be used to isolate novel alleles from relatives of crop species by positional cloning strategies.

The various types of maps discussed above can be used with the SDFs of the invention to identify Quantitative Trait Loci (QTLs). Many important crop traits, such as the solids content of tomatoes, are quantitative traits and result from the combined interactions of several genes. These genes reside at different loci in the genome, oftentimes on different chromosomes, and generally exhibit multiple alleles at each locus. The SDFs of the invention can be used to identify QTLs and isolate specific alleles as described by de Vicente and Tanksley (*Genetics* 134:585 (1993)). In addition to isolating QTL alleles in present crop species, the SDFs of the invention can also be used to isolate alleles from the corresponding QTL of wild relatives. Transgenic plants having various combinations of QTL alleles can then be created and the effects of the combinations measured. Once a desired allele combination has been identified, crop improvement can be accomplished either through biotechnological means or by directed

conventional breeding programs (for review *see* Tanksley and McCouch, *Science* 277:1063 (1997)).

In another embodiment, the SDFs can be used to help create physical maps of the genome of corn, *Arabidopsis* and related species. Where SDFs have been ordered on a genetic map, as described above, they can be used as probes to discover which clones in large libraries of plant DNA fragments in YACs, BACs, etc. contain the same SDF or similar sequences, thereby facilitating the assignment of the large DNA fragments to chromosomal positions. Subsequently, the large BACs, YACs, etc. can be ordered unambiguously by more detailed studies of their sequence composition (e.g. Marra et al. (1997) *Genomic Research* 7:1072-1084) and by using their end or other sequences to find the identical sequences in other cloned DNA fragments. The overlapping of DNA sequences in this way allows large contigs of plant sequences to be built that, when sufficiently extended, provide a complete physical map of a chromosome. Sometimes the SDFs themselves will provide the means of joining cloned sequences into a contig.

The patent publication WO95/35505 and U.S. Patents 5,445,943 and 5,410,270 describe scanning multiple alleles of a plurality of loci using hybridization to arrays of oligonucleotides. These techniques are useful for each of the types of mapping discussed above.

Following the procedures described above and using a plurality of the SDFs of the present invention, any individual can be genotyped. These individual genotypes can be used for the identification of particular cultivars, varieties, lines, ecotypes and genetically modified plants or can serve as tools for subsequent genetic studies involving multiple phenotypic traits.

### B.3 Southern Blot Hybridization

The sequences from Reference and Sequence and those encoding polypeptides of Protein Group and Protein Group Matrix tables or fragments thereof can be used as probes for various hybridization techniques. These techniques are useful for detecting target polynucleotides in a sample or for determining whether transgenic plants, seeds or host cells harbor a gene or sequence of interest and thus might be expected to exhibit a particular trait or phenotype.

In addition, the SDFs from the invention can be used to isolate additional members of gene families from the same or different species and/or orthologous genes from the same or different species. This is accomplished by hybridizing an SDF to, for example, a Southern blot containing the appropriate genomic DNA or cDNA. Given the resulting hybridization data, one

of ordinary skill in the art could distinguish and isolate the correct DNA fragments by size, restriction sites, sequence and stated hybridization conditions from a gel or from a library.

Identification and isolation of orthologous genes from closely related species and alleles within a species is particularly desirable because of their potential for crop improvement. Many important crop traits, such as the solid content of tomatoes, result from the combined interactions of the products of several genes residing at different loci in the genome. Generally, alleles at each of these loci can make quantitative differences to the trait. By identifying and isolating numerous alleles for each locus from within or different species, transgenic plants with various combinations of alleles can be created and the effects of the combinations measured. Once a more favorable allele combination has been identified, crop improvement can be accomplished either through biotechnological means or by directed conventional breeding programs (Tanksley et al. *Science* 277:1063(1997)).

The results from hybridizations of the SDFs of the invention to, for example, Southern blots containing DNA from another species can also be used to generate restriction fragment maps for the corresponding genomic regions. These maps provide additional information about the relative positions of restriction sites within fragments, further distinguishing mapped DNA from the remainder of the genome.

Physical maps can be made by digesting genomic DNA with different combinations of restriction enzymes.

Probes for Southern blotting to distinguish individual restriction fragments can range in size from 15 to 20 nucleotides to several thousand nucleotides. More preferably, the probe is 100 to 1,000 nucleotides long for identifying members of a gene family when it is found that repetitive sequences would complicate the hybridization. For identifying an entire corresponding gene in another species, the probe is more preferably the length of the gene, typically 2,000 to 10,000 nucleotides, but probes 50-1,000 nucleotides long might be used. Some genes, however, might require probes up to 1,500 nucleotides long or overlapping probes constituting the full-length sequence to span their lengths.

Also, while it is preferred that the probe be homogeneous with respect to its sequence, it is not necessary. For example, as described below, a probe representing members of a gene family having diverse sequences can be generated using PCR to amplify genomic DNA or RNA templates using primers derived from SDFs that include sequences that define the gene family.

For identifying corresponding genes in another species, the next most preferable probe is a cDNA spanning the entire coding sequence, which allows all of the mRNA-coding fragment of the gene to be identified. Probes for Southern blotting can easily be generated from SDFs by making primers having the sequence at the ends of the SDF and using corn or *Arabidopsis* genomic DNA as a template. In instances where the SDF includes sequence conserved among species, primers including the conserved sequence can be used for PCR with genomic DNA from a species of interest to obtain a probe.

Similarly, if the SDF includes a domain of interest, that fragment of the SDF can be used to make primers and, with appropriate template DNA, used to make a probe to identify genes containing the domain. Alternatively, the PCR products can be resolved, for example by gel electrophoresis, and cloned and/or sequenced. Using Southern hybridization, the variants of the domain among members of a gene family, both within and across species, can be examined.

#### B.4.1 Isolating DNA from Related Organisms

The SDFs of the invention can be used to isolate the corresponding DNA from other organisms. Either cDNA or genomic DNA can be isolated. For isolating genomic DNA, a lambda, cosmid, BAC or YAC, or other large insert genomic library from the plant of interest can be constructed using standard molecular biology techniques as described in detail by Sambrook et al. 1989 (Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, New York) and by Ausubel et al. 1992 (Current Protocols in Molecular Biology, Greene Publishing, New York).

To screen a phage library, for example, recombinant lambda clones are plated out on appropriate bacterial medium using an appropriate *E. coli* host strain. The resulting plaques are lifted from the plates using nylon or nitrocellulose filters. The plaque lifts are processed through denaturation, neutralization, and washing treatments following the standard protocols outlined by Ausubel et al. (1992). The plaque lifts are hybridized to either radioactively labeled or non-radioactively labeled SDF DNA at room temperature for about 16 hours, usually in the presence of 50% formamide and 5X SSC (sodium chloride and sodium citrate) buffer and blocking reagents. The plaque lifts are then washed at 42°C with 1% Sodium Dodecyl Sulfate (SDS) and at a particular concentration of SSC. The SSC concentration used is dependent upon the stringency at which hybridization occurred in the initial Southern blot analysis performed. For



example, if a fragment hybridized under medium stringency (e.g.,  $T_m - 20^\circ\text{C}$ ), then this condition is maintained or preferably adjusted to a less stringent condition (e.g.,  $T_m - 30^\circ\text{C}$ ) to wash the plaque lifts. Positive clones show detectable hybridization e.g., by exposure to X-ray films or chromogen formation. The positive clones are then subsequently isolated for  
5 purification using the same general protocol outlined above. Once the clone is purified, restriction analysis can be conducted to narrow the region corresponding to the gene of interest. The restriction analysis and succeeding subcloning steps can be done using procedures described by, for example Sambrook et al. (1989) cited above.

The procedures outlined for the lambda library are essentially similar to those used for YAC library screening, except that the YAC clones are harbored in bacterial colonies. The YAC clones are plated out at reasonable density on nitrocellulose or nylon filters supported by appropriate bacterial medium in petri plates. Following the growth of the bacterial clones, the filters are processed through the denaturation, neutralization, and washing steps following the procedures of Ausubel et al. 1992. The same hybridization procedures for lambda library  
10 screening are followed.

To isolate cDNA, similar procedures using appropriately modified vectors are employed. For instance, the library can be constructed in a lambda vector appropriate for cloning cDNA such as  $\lambda\text{gt}11$ . Alternatively, the cDNA library can be made in a plasmid vector. cDNA for cloning can be prepared by any of the methods known in the art, but is preferably prepared as  
15 described above. Preferably, a cDNA library will include a high proportion of full-length clones.

#### B. 5. Isolating and/or Identifying Orthologous Genes

Probes and primers of the invention can be used to identify and/or isolate polynucleotides related to those in the Reference, Sequence, Protein Group, and Protein Group Matrix tables.

25 Related polynucleotides are those that are native to other plant organisms and exhibit either similar sequence or encode polypeptides with similar biological activity. One specific example is an orthologous gene. Orthologous genes have the same functional activity. As such, orthologous genes may be distinguished from homologous genes. The percentage of identity is a function of evolutionary separation and, in closely related species, the percentage of identity can be 98 to  
30 100%. The amino acid sequence of a protein encoded by an orthologous gene can be less than 75%

identical, but tends to be at least 75% or at least 80% identical, more preferably at least 90%, most preferably at least 95% identical to the amino acid sequence of the reference protein.

To find orthologous genes, the probes are hybridized to nucleic acids from a species of interest under low stringency conditions, preferably one where sequences containing as much as 40-45% mismatches will be able to hybridize. This condition is established by  $T_m - 40^\circ\text{C}$  to  $T_m - 48^\circ\text{C}$  (see below). Blots are then washed under conditions of increasing stringency. It is preferable that the wash stringency be such that sequences that are 85 to 100% identical will hybridize. More preferably, sequences 90 to 100% identical will hybridize and most preferably only sequences greater than 95% identical will hybridize. One of ordinary skill in the art will recognize that, due to degeneracy in the genetic code, amino acid sequences that are identical can be encoded by DNA sequences as little as 67% identical or less. Thus, it is preferable, for example, to make an overlapping series of shorter probes, on the order of 24 to 45 nucleotides, and individually hybridize them to the same arrayed library to avoid the problem of degeneracy introducing large numbers of mismatches.

As evolutionary divergence increases, genome sequences also tend to diverge. Thus, one of skill will recognize that searches for orthologous genes between more divergent species will require the use of lower stringency conditions compared to searches between closely related species. Also, degeneracy of the genetic code is more of a problem for searches in the genome of a species more distant evolutionarily from the species that is the source of the SDF probe sequences.

Therefore the method described in Bouckaert et al., U.S. Ser. No. 60/121,700 Atty. Dkt. No. 2750-117P, Client Dkt. No. 00010.001, filed February 25, 1999, hereby incorporated in its entirety by reference, can be applied to the SDFs of the present invention to isolate related genes from plant species which do not hybridize to the corn *Arabidopsis*, soybean, rice, wheat, and other plant sequences of the reference, Sequence, Protein Group, and Protein Group Matrix tables.

Identification of the relationship of nucleotide or amino acid sequences among plant species can be done by comparing the nucleotide or amino acid sequences of SDFs of the present application with nucleotide or amino acid sequences of other SDFs such as those present in applications listed in the table below:

The SDFs of the invention can also be used as probes to search for genes that are related to the SDF within a species. Such related genes are typically considered to be members of a gene family. In such a case, the sequence similarity will often be concentrated into one or a few fragments of the sequence. The fragments of similar sequence that define the gene family typically encode a fragment of a protein or RNA that has an enzymatic or structural function. The percentage of identity in the amino acid sequence of the domain that defines the gene family is preferably at least 70%, more preferably 80 to 95%, most preferably 85 to 99%. To search for members of a gene family within a species, a low stringency hybridization is usually performed, but this will depend upon the size, distribution and degree of sequence divergence of domains that define the gene family. SDFs encompassing regulatory regions can be used to identify coordinately expressed genes by using the regulatory region sequence of the SDF as a probe.

In the instances where the SDFs are identified as being expressed from genes that confer a particular phenotype, then the SDFs can also be used as probes to assay plants of different species for those phenotypes.

#### I.C. Methods to Inhibit Gene Expression

The nucleic acid molecules of the present invention can be used to inhibit gene transcription and/or translation. Example of such methods include, without limitation:

Antisense Constructs;  
Ribozyme Constructs;  
Chimeraplast Constructs;  
Co-Suppression;  
Transcriptional Silencing; and  
Other Methods of Gene Expression.

#### C.1 Antisense

In some instances it is desirable to suppress expression of an endogenous or exogenous gene. A well-known instance is the FLAVOR-SAVOR™ tomato, in which the gene encoding ACC synthase is inactivated by an antisense approach, thus delaying softening of the fruit after ripening. See for example, U.S. Patent No. 5,859,330; U.S. Patent No. 5,723,766; Oeller, et al,

Science, 254:437-439(1991); and Hamilton et al, *Nature*, 346:284-287 (1990). Also, timing of flowering can be controlled by suppression of the *FLOWERING LOCUS C (FLC)*; high levels of this transcript are associated with late flowering, while absence of *FLC* is associated with early flowering (S.D. Michaels et al., *Plant Cell* 11:949 (1999). Also, the transition of apical meristem from production of leaves with associated shoots to flowering is regulated by *TERMINAL FLOWER1*, *APETALA1* and *LEAFY*. Thus, when it is desired to induce a transition from shoot production to flowering, it is desirable to suppress *TFL1* expression (S.J. Liljegren, *Plant Cell* 11:1007 (1999)). As another instance, arrested ovule development and female sterility result from suppression of the ethylene forming enzyme but can be reversed by application of ethylene (D. De Martinis et al., *Plant Cell* 11:1061 (1999)). The ability to manipulate female fertility of plants is useful in increasing fruit production and creating hybrids.

In the case of polynucleotides used to inhibit expression of an endogenous gene, the introduced sequence need not be perfectly identical to a sequence of the target endogenous gene. The introduced polynucleotide sequence will typically be at least substantially identical to the target endogenous sequence.

Some polynucleotide SDFs in the Reference, Sequence, Protein Group, and Protein Group Matrix tables represent sequences that are expressed in corn, wheat, rice, soybean *Arabidopsis* and/or other plants. Thus the invention includes using these sequences to generate antisense constructs to inhibit translation and/or degradation of transcripts of said SDFs, typically in a plant cell.

To accomplish this, a polynucleotide segment from the desired gene that can hybridize to the mRNA expressed from the desired gene (the "antisense segment") is operably linked to a promoter such that the antisense strand of RNA will be transcribed when the construct is present in a host cell. A regulated promoter can be used in the construct to control transcription of the antisense segment so that transcription occurs only under desired circumstances.

The antisense segment to be introduced generally will be substantially identical to at least a fragment of the endogenous gene or genes to be repressed. The sequence, however, need not be perfectly identical to inhibit expression. Further, the antisense product may hybridize to the untranslated region instead of or in addition to the coding sequence of the gene. The vectors of the present invention can be designed such that the inhibitory effect applies to other proteins within a family of genes exhibiting homology or substantial homology to the target gene.

For antisense suppression, the introduced antisense segment sequence also need not be full length relative to either the primary transcription product or the fully processed mRNA. Generally, a higher percentage of sequence identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments may be equally effective. Normally, a sequence of between about 30 or 40 nucleotides and the full length of the transcript can be used, though a sequence of at least about 100 nucleotides is preferred, a sequence of at least about 200 nucleotides is more preferred, and a sequence of at least about 500 nucleotides is especially preferred.

## C.2. Ribozymes

It is also contemplated that gene constructs representing ribozymes and based on the SDFs in the Reference and Sequence tables or those encoding polypeptides of the Protein Group and Protein Group Matrix tables and fragment thereof are an object of the invention. Ribozymes can also be used to inhibit expression of genes by suppressing the translation of the mRNA into a polypeptide. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs.

A number of classes of ribozymes have been identified. One class of ribozymes is derived from a number of small circular RNAs, which are capable of self-cleavage and replication in plants. The RNAs replicate either alone (viroid RNAs) or with a helper virus (satellite RNAs). Examples include RNAs from avocado sunblotch viroid and the satellite RNAs from tobacco ringspot virus, lucerne transient streak virus, velvet tobacco mottle virus, solanum nodiflorum mottle virus and subterranean clover mottle virus. The design and use of target RNA-specific ribozymes is described in Haseloff et al. *Nature*, 334:585 (1988).

Like the antisense constructs above, the ribozyme sequence fragment necessary for pairing need not be identical to the target nucleotides to be cleaved, nor identical to the sequences in the Reference and Sequence tables or those encoding polypeptide of the Protein Group and Protein Group Matrix tables or fragments thereof. Ribozymes may be constructed by

combining the ribozyme sequence and some fragment of the target gene which would allow recognition of the target gene mRNA by the resulting ribozyme molecule. Generally, the sequence in the ribozyme capable of binding to the target sequence exhibits a percentage of sequence identity with at least 80%, preferably with at least 85%, more preferably with at least 90% and most preferably with at least 95%, even more preferably, with at least 96%, 97%, 98% or 99% sequence identity to some fragment of a sequence in the Reference, Sequence, Protein Group, and Protein Group Matrix tables or the complement thereof. The ribozyme can be equally effective in inhibiting mRNA translation by cleaving either in the untranslated or coding regions. Generally, a higher percentage of sequence identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments may be equally effective.

### C.3. Chimeraplasts

The SDFs of the invention, such as those described by Reference, Sequence, Protein Group, and Protein Group Matrix tables, can also be used to construct chimeraplasts that can be introduced into a cell to produce at least one specific nucleotide change in a sequence corresponding to the SDF of the invention. A chimeraplast is an oligonucleotide comprising DNA and/or RNA that specifically hybridizes to a target region in a manner which creates a mismatched base-pair. This mismatched base-pair signals the cell's repair enzyme machinery which acts on the mismatched region resulting in the replacement, insertion or deletion of designated nucleotide(s). The altered sequence is then expressed by the cell's normal cellular mechanisms. Chimeraplasts can be designed to repair mutant genes, modify genes, introduce site-specific mutations, and/or act to interrupt or alter normal gene function (US Pat. Nos. 6,010,907 and 6,004,804; and PCT Pub. No. WO99/58723 and WO99/07865).

### C.4. Sense Suppression

The SDFs of the reference, Sequence, Protein Group, and Protein Group Matrix tables of the present invention are also useful to modulate gene expression by sense suppression. Sense suppression represents another method of gene suppression by introducing at least one exogenous copy or fragment of the endogenous sequence to be suppressed.

Introduction of expression cassettes in which a nucleic acid is configured in the sense orientation with respect to the promoter into the chromosome of a plant or by a self-replicating virus has been shown to be an effective means by which to induce degradation of mRNAs of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279 (1990), and U.S. Patents Nos. 5,034,323, 5,231,020, and 5,283,184. Inhibition of expression may require some transcription of the introduced sequence.

For sense suppression, the introduced sequence generally will be substantially identical to the endogenous sequence intended to be inactivated. The minimal percentage of sequence identity will typically be greater than about 65%, but a higher percentage of sequence identity might exert a more effective reduction in the level of normal gene products. Sequence identity of more than about 80% is preferred, though about 95% to absolute identity would be most preferred. As with antisense regulation, the effect would likely apply to any other proteins within a similar family of genes exhibiting homology or substantial homology to the suppressing sequence.

#### C.5. Transcriptional Silencing

The nucleic acid sequences of the invention, including the SDFs of the reference, Sequence, Protein Group, and Protein Group Matrix tables, and fragments thereof, contain sequences that can be inserted into the genome of an organism resulting in transcriptional silencing. Such regulatory sequences need not be operatively linked to coding sequences to modulate transcription of a gene. Specifically, a promoter sequence without any other element of a gene can be introduced into a genome to transcriptionally silence an endogenous gene (see, for example, Vaucheret, H et al. (1998) *The Plant Journal* 16: 651-659). As another example, triple helices can be formed using oligonucleotides based on sequences from Reference, Sequence, Protein Group, and Protein Group Matrix tables, fragments thereof, and substantially similar sequence thereto. The oligonucleotide can be delivered to the host cell and can bind to the promoter in the genome to form a triple helix and prevent transcription. An oligonucleotide of interest is one that can bind to the promoter and block binding of a transcription factor to the promoter. In such a case, the oligonucleotide can be complementary to the sequences of the promoter that interact with transcription binding factors.

#### C.6. Other Methods to Inhibit Gene Expression

Yet another means of suppressing gene expression is to insert a polynucleotide into the gene of interest to disrupt transcription or translation of the gene.

Low frequency homologous recombination can be used to target a polynucleotide insert to a gene by flanking the polynucleotide insert with sequences that are substantially similar to the gene to be disrupted. Sequences from Reference, Sequence, Protein Group, and Protein Group Matrix tables, fragments thereof, and substantially similar sequence thereto can be used for homologous recombination.

In addition, random insertion of polynucleotides into a host cell genome can also be used to disrupt the gene of interest. Azpiroz-Leehan et al., *Trends in Genetics* 13:152 (1997). In this method, screening for clones from a library containing random insertions is preferred to identifying those that have polynucleotides inserted into the gene of interest. Such screening can be performed using probes and/or primers described above based on sequences from Reference, Sequence, Protein Group, and Protein Group Matrix tables, fragments thereof, and substantially similar sequence thereto. The screening can also be performed by selecting clones or R<sub>1</sub> plants having a desired phenotype.

#### I.D. Methods of Functional Analysis

The constructs described in the methods under I.C. above can be used to determine the function of the polypeptide encoded by the gene that is targeted by the constructs.

Down-regulating the transcription and translation of the targeted gene in the host cell or organisms, such as a plant, may produce phenotypic changes as compared to a wild-type cell or organism. In addition, *in vitro* assays can be used to determine if any biological activity, such as calcium flux, DNA transcription, nucleotide incorporation, etc., are being modulated by the down-regulation of the targeted gene.

Coordinated regulation of sets of genes, e.g., those contributing to a desired polygenic trait, is sometimes necessary to obtain a desired phenotype. SDFs of the invention representing transcription activation and DNA binding domains can be assembled into hybrid transcriptional activators. These hybrid transcriptional activators can be used with their corresponding DNA elements (i.e., those bound by the DNA-binding SDFs) to effect coordinated expression of desired genes (J.J. Schwarz et al., *Mol. Cell. Biol.* 12:266 (1992), A. Martinez et al., *Mol. Gen. Genet.* 261:546 (1999)).



The SDFs of the invention can also be used in the two-hybrid genetic systems to identify networks of protein-protein interactions (L. McAlister-Henn et al., *Methods* 19:330 (1999), J.C. Hu et al., *Methods* 20:80 (2000), M. Golovkin et al., *J. Biol. Chem.* 274:36428 (1999), K. Ichimura et al., *Biochem. Biophys. Res. Comm.* 253:532 (1998)). The SDFs of the invention can also be used in various expression display methods to identify important protein-DNA interactions (e.g. B. Luo et al., *J. Mol. Biol.* 266:479 (1997)).

#### I.E. Promoters

The SDFs of the invention are also useful as structural or regulatory sequences in a construct for modulating the expression of the corresponding gene in a plant or other organism, e.g. a symbiotic bacterium. For example, promoter sequences associated to SDFs of the reference, Sequence, Protein Group, and Protein Group Matrix tables of the present invention can be useful in directing expression of coding sequences either as constitutive promoters or to direct expression in particular cell types, tissues, or organs or in response to environmental stimuli.

With respect to the SDFs of the present invention a promoter is likely to be a relatively small portion of a genomic DNA (gDNA) sequence located in the first 2000 nucleotides upstream from an initial exon identified in a gDNA sequence or initial "ATG" or methionine codon or translational start site in a corresponding cDNA sequence. Such promoters are more likely to be found in the first 1000 nucleotides upstream of an initial ATG or methionine codon or translational start site of a cDNA sequence corresponding to a gDNA sequence. In particular, the promoter is usually located upstream of the transcription start site. The fragments of a particular gDNA sequence that function as elements of a promoter in a plant cell will preferably be found to hybridize to gDNA sequences presented and described in the Reference table at medium or high stringency, relevant to the length of the probe and its base composition.

Promoters are generally modular in nature. Promoters can consist of a basal promoter that functions as a site for assembly of a transcription complex comprising an RNA polymerase, for example RNA polymerase II. A typical transcription complex will include additional factors such as TF<sub>II</sub>B, TF<sub>II</sub>D, and TF<sub>II</sub>E. Of these, TF<sub>II</sub>D appears to be the only one to bind DNA directly. The promoter might also contain one or more enhancers and/or suppressors that function as binding sites for additional transcription factors that have the function of modulating the level of transcription

with respect to tissue specificity and of transcriptional responses to particular environmental or nutritional factors, and the like.

Short DNA sequences representing binding sites for proteins can be separated from each other by intervening sequences of varying length. For example, within a particular functional module, protein binding sites may be constituted by regions of 5 to 60, preferably 10 to 30, more preferably 10 to 20 nucleotides. Within such binding sites, there are typically 2 to 6 nucleotides that specifically contact amino acids of the nucleic acid binding protein. The protein binding sites are usually separated from each other by 10 to several hundred nucleotides, typically by 15 to 150 nucleotides, often by 20 to 50 nucleotides. DNA binding sites in promoter elements often display dyad symmetry in their sequence. Often elements binding several different proteins, and/or a plurality of sites that bind the same protein, will be combined in a region of 50 to 1,000 basepairs.

Elements that have transcription regulatory function can be isolated from their corresponding endogenous gene, or the desired sequence can be synthesized, and recombined in constructs to direct expression of a coding region of a gene in a desired tissue-specific, temporal-specific or other desired manner of inducibility or suppression. When hybridizations are performed to identify or isolate elements of a promoter by hybridization to the long sequences presented in the Reference tables, conditions are adjusted to account for the above-described nature of promoters. For example short probes, constituting the element sought, are preferably used under low temperature and/or high salt conditions. When long probes, which might include several promoter elements are used, low to medium stringency conditions are preferred when hybridizing to promoters across species.

If a nucleotide sequence of an SDF, or part of the SDF, functions as a promoter or fragment of a promoter, then nucleotide substitutions, insertions or deletions that do not substantially affect the binding of relevant DNA binding proteins would be considered equivalent to the exemplified nucleotide sequence. It is envisioned that there are instances where it is desirable to decrease the binding of relevant DNA binding proteins to silence or down-regulate a promoter, or conversely to increase the binding of relevant DNA binding proteins to enhance or up-regulate a promoter and vice versa. In such instances, polynucleotides representing changes to the nucleotide sequence of the DNA-protein contact region by insertion of additional nucleotides, changes to identity of relevant nucleotides, including use of chemically-modified bases, or deletion of one or more nucleotides are considered encompassed

by the present invention. In addition, fragments of the promoter sequences described by Reference tables and variants thereof can be fused with other promoters or fragments to facilitate transcription and/or transcription in specific type of cells or under specific conditions.

Promoter function can be assayed by methods known in the art, preferably by measuring activity of a reporter gene operatively linked to the sequence being tested for promoter function. Examples of reporter genes include those encoding luciferase, green fluorescent protein, GUS, neo, cat and bar.

#### I.F. UTRs and Junctions

Polynucleotides comprising untranslated (UTR) sequences and intron/exon junctions are also within the scope of the invention. UTR sequences include introns and 5' or 3' untranslated regions ( 5' UTRs or 3' UTRs). Fragments of the sequences shown in the Reference and Sequence tables can comprise UTRs and intron/exon junctions.

These fragments of SDFs, especially UTRs, can have regulatory functions related to, for example, translation rate and mRNA stability. Thus, these fragments of SDFs can be isolated for use as elements of gene constructs for regulated production of polynucleotides encoding desired polypeptides.

Introns of genomic DNA segments might also have regulatory functions. Sometimes regulatory elements, especially transcription enhancer or suppressor elements, are found within introns. Also, elements related to stability of heteronuclear RNA and efficiency of splicing and of transport to the cytoplasm for translation can be found in intron elements. Thus, these segments can also find use as elements of expression vectors intended for use to transform plants.

Just as with promoters UTR sequences and intron/exon junctions can vary from those shown in the Reference and Sequence tables. Such changes from those sequences preferably will not affect the regulatory activity of the UTRs or intron/exon junction sequences on expression, transcription, or translation unless selected to do so. However, in some instances, down- or up-regulation of such activity may be desired to modulate traits or phenotypic or *in vitro* activity.

#### I.G. Coding Sequences

Isolated polynucleotides of the invention can include coding sequences that encode polypeptides comprising an amino acid sequence encoded by sequences described in the Reference and Sequence tables or an amino acid sequence presented in the Reference, Sequence, Protein Group, and Protein Group Matrix tables.

5 A nucleotide sequence encodes a polypeptide if a cell (or a cell free *in vitro* system) expressing that nucleotide sequence produces a polypeptide having the recited amino acid sequence when the nucleotide sequence is transcribed and the primary transcript is subsequently processed and translated by a host cell (or a cell free *in vitro* system) harboring the nucleic acid. Thus, an isolated nucleic acid that encodes a particular amino acid sequence can be a genomic  
10 sequence comprising exons and introns or a cDNA sequence that represents the product of splicing thereof. An isolated nucleic acid encoding an amino acid sequence also encompasses heteronuclear RNA, which contains sequences that are spliced out during expression, and mRNA, which lacks those sequences.

Coding sequences can be constructed using chemical synthesis techniques or by isolating coding sequences or by modifying such synthesized or isolated coding sequences as described  
15 above.

In addition to coding sequences encoding the polypeptide sequences of the reference, Sequence, Protein Group, and Protein Group Matrix tables, which are native to corn, *Arabidopsis*, soybean, rice, wheat, and other plants, the isolated polynucleotides can be  
20 polynucleotides that encode variants, fragments, and fusions of those native proteins. Such polypeptides are described below in part II.

In variant polynucleotides generally, the number of substitutions, deletions or insertions is preferably less than 20%, more preferably less than 15%; even more preferably less than 10%, 5%, 3% or 1% of the number of nucleotides comprising a particularly exemplified sequence. It is  
25 generally expected that non-degenerate nucleotide sequence changes that result in 1 to 10, more preferably 1 to 5 and most preferably 1 to 3 amino acid insertions, deletions or substitutions will not greatly affect the function of an encoded polypeptide. The most preferred embodiments are those wherein 1 to 20, preferably 1 to 10, most preferably 1 to 5 nucleotides are added to, or deleted from  
30 polynucleotides that encode polypeptides of the Protein Group, and Protein Group Matrix tables or fragments thereof.

Insertions or deletions in polynucleotides intended to be used for encoding a polypeptide preferably preserve the reading frame. This consideration is not so important in instances when the polynucleotide is intended to be used as a hybridization probe.

## II. Polypeptides and Proteins

### IIA. Native polypeptides and proteins

Polypeptides within the scope of the invention include both native proteins as well as variants, fragments, and fusions thereof. Polypeptides of the invention are those encoded by any of the six reading frames of sequences shown in the Reference and Sequence tables, preferably encoded by the three frames reading in the 5' to 3' direction of the sequences as shown.

Native polypeptides include the proteins encoded by the sequences shown in the Reference and Sequence tables. Such native polypeptides include those encoded by allelic variants.

Polypeptide and protein variants will exhibit at least 75% sequence identity to those native polypeptides of the Reference and Sequence tables. More preferably, the polypeptide variants will exhibit at least 85% sequence identity; even more preferably, at least 90% sequence identity; more preferably at least 95%, 96%, 97%, 98%, or 99% sequence identity. Fragments of polypeptide or fragments of polypeptides will exhibit similar percentages of sequence identity to the relevant fragments of the native polypeptide. Fusions will exhibit a similar percentage of sequence identity in that fragment of the fusion represented by the variant of the native peptide.

Polypeptide and protein variants of the invention will exhibit at least 75% sequence identity to those motifs or consensus sequences of the Protein Group and Protein Group Matrix tables. More preferably, the polypeptide variants will exhibit at least 85% sequence identity; even more preferably, at least 90% sequence identity; more preferably at least 95%, 96%, 97%, 98%, or 99% sequence identity. Fragments of polypeptide or fragments of polypeptides will exhibit similar percentages of sequence identity to the relevant fragments of the native polypeptide that are indicated in the Protein Group table. Fusions will exhibit a similar percentage of sequence identity in that fragment of the fusion represented by the variant of the native peptide.

Furthermore, polypeptide variants will exhibit at least one of the functional properties of the native protein. Such properties include, without limitation, protein interaction, DNA interaction,

biological activity, immunological activity, receptor binding, signal transduction, transcription activity, growth factor activity, secondary structure, three-dimensional structure, etc. As to properties related to *in vitro* or *in vivo* activities, the variants preferably exhibit at least 60% of the activity of the native protein; more preferably at least 70%, even more preferably at least 80%, 85%, 90% or 95% of at least one activity of the native protein.

One type of variant of native polypeptides comprises amino acid substitutions, deletions and/or insertions. Conservative substitutions are preferred to maintain the function or activity of the polypeptide.

Within the scope of percentage of sequence identity described above, a polypeptide of the invention may have additional individual amino acids or amino acid sequences inserted into the polypeptide in the middle thereof and/or at the N-terminal and/or C-terminal ends thereof. Likewise, some of the amino acids or amino acid sequences may be deleted from the polypeptide.

#### A.1 Antibodies

Isolated polypeptides can be utilized to produce antibodies. Polypeptides of the invention can generally be used, for example, as antigens for raising antibodies by known techniques. The resulting antibodies are useful as reagents for determining the distribution of the antigen protein within the tissues of a plant or within a cell of a plant. The antibodies are also useful for examining the production level of proteins in various tissues, for example in a wild-type plant or following genetic manipulation of a plant, by methods such as Western blotting.

Antibodies of the present invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the polypeptides of the invention are first used to immunize a suitable animal, such as a mouse, rat, rabbit, or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies as detection reagents. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization

using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization.

Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating the blood at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000xg for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the method of Kohler and Milstein, *Nature* 256: 495 (1975), or modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells can be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate, or well, coated with the protein antigen. B-cells producing membrane-bound immunoglobulin specific for the antigen bind to the plate, and are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected Mab-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

Other methods for sustaining antibody-producing B-cell clones, such as by EBV transformation, are known.

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly <sup>32</sup>P and <sup>125</sup>I), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TNB) to a blue pigment, quantifiable with a spectrophotometer.

## A.2 In Vitro Applications of Polypeptides

Some polypeptides of the invention will have enzymatic activities that are useful *in vitro*. For example, the soybean trypsin inhibitor (Kunitz) family is one of the numerous families of

proteainase inhibitors. It comprises plant proteins which have inhibitory activity against serine proteinases from the trypsin and subtilisin families, thiol proteinases and aspartic proteinases. Thus, these peptides find *in vitro* use in protein purification protocols and perhaps in therapeutic settings requiring topical application of protease inhibitors.

Delta-aminolevulinic acid dehydratase (EC 4.2.1.24) (ALAD) catalyzes the second step in the biosynthesis of heme, the condensation of two molecules of 5-aminolevulinate to form porphobilinogen and is also involved in chlorophyll biosynthesis(Kaczor et al. (1994) Plant Physiol. 1-4: 1411-7; Smith (1988) Biochem. J. 249: 423-8; Schneider (1976) Z. naturforsch. [C] 31: 55-63). Thus, ALAD proteins can be used as catalysts in synthesis of heme derivatives. Enzymes of biosynthetic pathways generally can be used as catalysts for *in vitro* synthesis of the compounds representing products of the pathway.

Polypeptides encoded by SDFs of the invention can be engineered to provide purification reagents to identify and purify additional polypeptides that bind to them. This allows one to identify proteins that function as multimers or elucidate signal transduction or metabolic pathways. In the case of DNA binding proteins, the polypeptide can be used in a similar manner to identify the DNA determinants of specific binding (S. Pierrou et al., *Anal. Biochem.* 229:99 (1995), S. Chusacutanachai et al., *J. Biol. Chem.* 274:23591 (1999), Q. Lin et al., *J. Biol. Chem.* 272:27274 (1997)).

## II.B. POLYPEPTIDE VARIANTS, FRAGMENTS, AND FUSIONS

Generally, variants, fragments, or fusions of the polypeptides encoded by the maximum length sequence(MLS) can exhibit at least one of the activities of the identified domains and/or related polypeptides described in Sections (C) and (D) of The Reference tables corresponding to the MLS of interest.

### II.B.(1) Variants

A type of variant of the native polypeptides comprises amino acid substitutions. Conservative substitutions, described above (see II.), are preferred to maintain the function or activity of the polypeptide. Such substitutions include conservation of charge, polarity, hydrophobicity, size, etc. For example, one or more amino acid residues within the sequence can be substituted with another amino acid of similar polarity that acts as a functional equivalent, for



example providing a hydrogen bond in an enzymatic catalysis. Substitutes for an amino acid within an exemplified sequence are preferably made among the members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Within the scope of percentage of sequence identity described above, a polypeptide of the invention may have additional individual amino acids or amino acid sequences inserted into the polypeptide in the middle thereof and/or at the N-terminal and/or C-terminal ends thereof. Likewise, some of the amino acids or amino acid sequences may be deleted from the polypeptide. Amino acid substitutions may also be made in the sequences; conservative substitutions being preferred.

One preferred class of variants are those that comprise (1) the domain of an encoded polypeptide and/or (2) residues conserved between the encoded polypeptide and related polypeptides. For this class of variants, the encoded polypeptide sequence is changed by insertion, deletion, or substitution at positions flanking the domain and/or conserved residues.

Another class of variants includes those that comprise an encoded polypeptide sequence that is changed in the domain or conserved residues by a conservative substitution.

Yet another class of variants includes those that lack one of the *in vitro* activities, or structural features of the encoded polypeptides. One example is polypeptides or proteins produced from genes comprising dominant negative mutations. Such a variant may comprise an encoded polypeptide sequence with non-conservative changes in a particular domain or group of conserved residues.

## II.A.(2) FRAGMENTS

Fragments of particular interest are those that comprise a domain identified for a polypeptide encoded by an MLS of the instant invention and variants thereof. Also, fragments that comprise at least one region of residues conserved between an MLS encoded polypeptide and its related polypeptides are of great interest. Fragments are sometimes useful as polypeptides corresponding to genes comprising dominant negative mutations are.

## II.A.(3) FUSIONS

Of interest are chimeras comprising (1) a fragment of the MLS encoded polypeptide or variants thereof of interest and (2) a fragment of a polypeptide comprising the same domain. For example, an AP2 helix encoded by a MLS of the invention fused to second AP2 helix from ANT protein, which comprises two AP2 helices. The present invention also encompasses fusions of  
5 MLS encoded polypeptides, variants, or fragments thereof fused with related proteins or fragments thereof.

### DEFINITION OF DOMAINS

10 The polypeptides of the invention may possess identifying domains as shown in The Reference tables. Specific domains within the MLS encoded polypeptides are indicated in The Reference tables. In addition, the domains within the MLS encoded polypeptide can be defined by the region that exhibits at least 70% sequence identity with the consensus sequences listed in the detailed description below of each of the domains.

15 The majority of the protein domain descriptions given in the protein domain table are obtained from Prosite, (<http://www.expasy.ch/prosite/>), and Pfam, (<http://pfam.wustl.edu/browse.shtml>). Examples of domain descriptions are listed in the Protein Domain table.

#### A. Activities of Polypeptides Comprising Signal Peptides

20 Polypeptides comprising signal peptides are a family of proteins that are typically targeted to (1) a particular organelle or intracellular compartment, (2) interact with a particular molecule or (3) for secretion outside of a host cell. Example of polypeptides comprising signal  
25 peptides include, without limitation, secreted proteins, soluble proteins, receptors, proteins retained in the ER, etc.

30 These proteins comprising signal peptides are useful to modulate ligand-receptor interactions, cell-to-cell communication, signal transduction, intracellular communication, and activities and/or chemical cascades that take part in an organism outside or within of any particular cell.

One class of such proteins are soluble proteins which are transported out of the cell. These proteins can act as ligands that bind to receptor to trigger signal transduction or to permit communication between cells.

Another class is receptor proteins which also comprise a retention domain that lodges the receptor protein in the membrane when the cell transports the receptor to the surface of the cell. Like the soluble ligands, receptors can also modulate signal transduction and communication between cells.

In addition the signal peptide itself can serve as a ligand for some receptors. An example is the interaction of the ER targeting signal peptide with the signal recognition particle (SRP). Here, the SRP binds to the signal peptide, halting translation, and the resulting SRP complex then binds to docking proteins located on the surface of the ER, prompting transfer of the protein into the ER.

A description of signal peptide residue composition is described below in Subsection IV.C.1.

### III. Methods of Modulating Polypeptide Production

It is contemplated that polynucleotides of the invention can be incorporated into a host cell or in-vitro system to modulate polypeptide production. For instance, the SDFs prepared as described herein can be used to prepare expression cassettes useful in a number of techniques for suppressing or enhancing expression.

An example are polynucleotides comprising sequences to be transcribed, such as coding sequences, of the present invention can be inserted into nucleic acid constructs to modulate polypeptide production. Typically, such sequences to be transcribed are heterologous to at least one element of the nucleic acid construct to generate a chimeric gene or construct.

Another example of useful polynucleotides are nucleic acid molecules comprising regulatory sequences of the present invention. Chimeric genes or constructs can be generated

when the regulatory sequences of the invention linked to heterologous sequences in a vector construct. Within the scope of invention are such chimeric gene and/or constructs.

Also within the scope of the invention are nucleic acid molecules, whereof at least a part or fragment of these DNA molecules are presented in the Reference and Sequence tables or polynucleotide encoding polypeptides of the Protein Group or Protein Group Matrix tables of the present application, and wherein the coding sequence is under the control of its own promoter and/or its own regulatory elements. Such molecules are useful for transforming the genome of a host cell or an organism regenerated from said host cell for modulating polypeptide production.

Additionally, a vector capable of producing the oligonucleotide can be inserted into the host cell to deliver the oligonucleotide.

More detailed description of components to be included in vector constructs are described both above and below.

Whether the chimeric vectors or native nucleic acids are utilized, such polynucleotides can be incorporated into a host cell to modulate polypeptide production. Native genes and/or nucleic acid molecules can be effective when exogenous to the host cell.

Methods of modulating polypeptide expression includes, without limitation:

Suppression methods, such as

Antisense

Ribozymes

Co-suppression

Insertion of Sequences into the Gene to be Modulated

Regulatory Sequence Modulation.

as well as Methods for Enhancing Production, such as

Insertion of Exogenous Sequences; and

Regulatory Sequence Modulation.

### III.A. Suppression

Expression cassettes of the invention can be used to suppress expression of endogenous genes which comprise the SDF sequence. Inhibiting expression can be useful, for instance, to

tailor the ripening characteristics of a fruit (Oeller et al., *Science* 254:437 (1991)) or to influence seed size (WO98/07842) or to provoke cell ablation (Mariani et al., *Nature* 357: 384-387 (1992)).

As described above, a number of methods can be used to inhibit gene expression in plants, such as antisense, ribozyme, introduction of exogenous genes into a host cell, insertion of a polynucleotide sequence into the coding sequence and/or the promoter of the endogenous gene of interest, and the like.

### III.A.1. Antisense

An expression cassette as described above can be transformed into host cell or plant to produce an antisense strand of RNA. For plant cells, antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, *see, e.g.*, Sheehy et al., *Proc. Nat. Acad. Sci. USA*, 85:8805 (1988), and Hiatt et al., U.S. Patent No. 4,801,340.

### III.A.2. Ribozymes

Similarly, ribozyme constructs can be transformed into a plant to cleave mRNA and down-regulate translation.

### III.A.3. Co-Suppression

Another method of suppression is by introducing an exogenous copy of the gene to be suppressed. Introduction of expression cassettes in which a nucleic acid is configured in the sense orientation with respect to the promoter has been shown to prevent the accumulation of mRNA. A detailed description of this method is described above.

### III.A.4. Insertion of Sequences into the Gene to be Modulated

Yet another means of suppressing gene expression is to insert a polynucleotide into the gene of interest to disrupt transcription or translation of the gene.

Homologous recombination could be used to target a polynucleotide insert to a gene using the Cre-Lox system (A.C. Vergunst et al., *Nucleic Acids Res.* 26:2729 (1998), A.C. Vergunst et al., *Plant Mol. Biol.* 38:393 (1998), H. Albert et al., *Plant J.* 7:649 (1995)).

In addition, random insertion of polynucleotides into a host cell genome can also be used to disrupt the gene of interest. Azpiroz-Leehan et al., *Trends in Genetics* 13:152 (1997). In this

method, screening for clones from a library containing random insertions is preferred for identifying those that have polynucleotides inserted into the gene of interest. Such screening can be performed using probes and/or primers described above based on sequences from the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables, fragments thereof, and substantially similar sequence thereto. The screening can also be performed by selecting clones or any transgenic plants having a desired phenotype.

#### III.A.5. Regulatory Sequence Modulation

The SDFs described in the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables, and fragments thereof are examples of nucleotides of the invention that contain regulatory sequences that can be used to suppress or inactivate transcription and/or translation from a gene of interest as discussed in I.C.5.

#### III.A.6. Genes Comprising Dominant-Negative Mutations

When suppression of production of the endogenous, native protein is desired it is often helpful to express a gene comprising a dominant negative mutation. Production of protein variants produced from genes comprising dominant negative mutations is a useful tool for research. Genes comprising dominant negative mutations can produce a variant polypeptide which is capable of competing with the native polypeptide, but which does not produce the native result. Consequently, over expression of genes comprising these mutations can titrate out an undesired activity of the native protein. For example, The product from a gene comprising a dominant negative mutation of a receptor can be used to constitutively activate or suppress a signal transduction cascade, allowing examination of the phenotype and thus the trait(s) controlled by that receptor and pathway. Alternatively, the protein arising from the gene comprising a dominant-negative mutation can be an inactive enzyme still capable of binding to the same substrate as the native protein and therefore competes with such native protein.

Products from genes comprising dominant-negative mutations can also act upon the native protein itself to prevent activity. For example, the native protein may be active only as a homomultimer or as one subunit of a heteromultimer. Incorporation of an inactive subunit into the multimer with native subunit(s) can inhibit activity.

Thus, gene function can be modulated in host cells of interest by insertion into these cells vector constructs comprising a gene comprising a dominant-negative mutation.

### III.B. Enhanced Expression

Enhanced expression of a gene of interest in a host cell can be accomplished by either (1) insertion of an exogenous gene; or (2) promoter modulation.

#### III.B.1. Insertion of an Exogenous Gene

Insertion of an expression construct encoding an exogenous gene can boost the number of gene copies expressed in a host cell.

Such expression constructs can comprise genes that either encode the native protein that is of interest or that encode a variant that exhibits enhanced activity as compared to the native protein. Such genes encoding proteins of interest can be constructed from the sequences from the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables, fragments thereof, and substantially similar sequence thereto.

Such an exogenous gene can include either a constitutive promoter permitting expression in any cell in a host organism or a promoter that directs transcription only in particular cells or times during a host cell life cycle or in response to environmental stimuli.

#### III.B.2. Regulatory Sequence Modulation

The SDFs of the Reference and Sequence tables, and fragments thereof, contain regulatory sequences that can be used to enhance expression of a gene of interest. For example, some of these sequences contain useful enhancer elements. In some cases, duplication of enhancer elements or insertion of exogenous enhancer elements will increase expression of a desired gene from a particular promoter. As other examples, all II promoters require binding of a regulatory protein to be activated, while some promoters may need a protein that signals a promoter binding protein to expose a polymerase binding site. In either case, over-production of such proteins can be used to enhance expression of a gene of interest by increasing the activation time of the promoter.

Such regulatory proteins are encoded by some of the sequences in the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables, fragments thereof, and substantially similar sequences thereto.

Coding sequences for these proteins can be constructed as described above.

#### IV. Gene Constructs and Vector Construction

To use isolated SDFs of the present invention or a combination of them or parts and/or mutants and/or fusions of said SDFs in the above techniques, recombinant DNA vectors which comprise said SDFs and are suitable for transformation of cells, such as plant cells, are usually prepared. The SDF construct can be made using standard recombinant DNA techniques (Sambrook et al. 1989) and can be introduced to the species of interest by *Agrobacterium*-mediated transformation or by other means of transformation (e.g., particle gun bombardment) as referenced below.

The vector backbone can be any of those typical in the art such as plasmids, viruses, artificial chromosomes, BACs, YACs and PACs and vectors of the sort described by

- (a) **BAC:** Shizuya et al., Proc. Natl. Acad. Sci. USA 89: 8794-8797 (1992); Hamilton et al., Proc. Natl. Acad. Sci. USA 93: 9975-9979 (1996);
- (b) **YAC:** Burke et al., Science 236:806-812 (1987);.
- (c) **PAC:** Sternberg N. et al., Proc Natl Acad Sci U S A. Jan;87(1):103-7 (1990);
- (d) **Bacteria-Yeast Shuttle Vectors:** Bradshaw et al., Nucl Acids Res 23: 4850-4856 (1995);
- (e) **Lambda Phage Vectors:** Replacement Vector, e.g., Frischauf et al., J. Mol Biol 170: 827-842 (1983); or Insertion vector, e.g., Huynh et al., In: Glover NM (ed) DNA Cloning: A practical Approach, Vol.1 Oxford: IRL Press (1985);
- (f) **T-DNA gene fusion vectors :** Walden et al., Mol Cell Biol 1: 175-194 (1990); and
- (g) **Plasmid vectors:** Sambrook et al., infra.

Typically, a vector will comprise the exogenous gene, which in its turn comprises an SDF of the present invention to be introduced into the genome of a host cell, and which gene may be an antisense construct, a ribozyme construct chimera, or a coding sequence with any desired transcriptional and/or translational regulatory sequences, such as promoters, UTRs, and 3' end termination sequences. Vectors of the invention can also include origins of replication, scaffold attachment regions (SARs), markers, homologous sequences, introns, etc.



A DNA sequence coding for the desired polypeptide, for example a cDNA sequence encoding a full length protein, will preferably be combined with transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the gene in the intended tissues of the transformed plant.

For example, for over-expression, a plant promoter fragment may be employed that will direct transcription of the gene in all tissues of a regenerated plant. Alternatively, the plant promoter may direct transcription of an SDF of the invention in a specific tissue (tissue-specific promoters) or may be otherwise under more precise environmental control (inducible promoters).

If proper polypeptide production is desired, a polyadenylation region at the 3'-end of the coding region is typically included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA.

The vector comprising the sequences from genes or SDF of the invention may comprise a marker gene that confers a selectable phenotype on plant cells. The vector can include promoter and coding sequence, for instance. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or phosphinotricin.

#### IV.A. Coding Sequences

Generally, the sequence in the transformation vector and to be introduced into the genome of the host cell does not need to be absolutely identical to an SDF of the present invention. Also, it is not necessary for it to be full length, relative to either the primary transcription product or fully processed mRNA. Furthermore, the introduced sequence need not have the same intron or exon pattern as a native gene. Also, heterologous non-coding segments can be incorporated into the coding sequence without changing the desired amino acid sequence of the polypeptide to be produced.

#### IV.B. Promoters

As explained above, introducing an exogenous SDF from the same species or an orthologous SDF from another species are useful to modulate the expression of a native gene corresponding to that SDF of interest. Such an SDF construct can be under the control of either a constitutive promoter or a highly regulated inducible promoter (e.g., a copper inducible

promoter). The promoter of interest can initially be either endogenous or heterologous to the species in question. When re-introduced into the genome of said species, such promoter becomes exogenous to said species. Over-expression of an SDF transgene can lead to co-suppression of the homologous endogeneous sequence thereby creating some alterations in the phenotypes of the transformed species as demonstrated by similar analysis of the chalcone synthase gene (Napoli et al., *Plant Cell* 2:279 (1990) and van der Krol et al., *Plant Cell* 2:291 (1990)). If an SDF is found to encode a protein with desirable characteristics, its over-production can be controlled so that its accumulation can be manipulated in an organ- or tissue-specific manner utilizing a promoter having such specificity.

Likewise, if the promoter of an SDF (or an SDF that includes a promoter) is found to be tissue-specific or developmentally regulated, such a promoter can be utilized to drive or facilitate the transcription of a specific gene of interest (e.g., seed storage protein or root-specific protein). Thus, the level of accumulation of a particular protein can be manipulated or its spatial localization in an organ- or tissue- specific manner can be altered.

#### IV. C Signal Peptides

SDFs of the present invention containing signal peptides are indicated in the Reference, Sequence, the Protein Group and Protein Group Matrix tables. In some cases it may be desirable for the protein encoded by an introduced exogenous or orthologous SDF to be targeted (1) to a particular organelle intracellular compartment, (2) to interact with a particular molecule such as a membrane molecule or (3) for secretion outside of the cell harboring the introduced SDF. This will be accomplished using a signal peptide.

Signal peptides direct protein targeting, are involved in ligand-receptor interactions and act in cell to cell communication. Many proteins, especially soluble proteins, contain a signal peptide that targets the protein to one of several different intracellular compartments. In plants, these compartments include, but are not limited to, the endoplasmic reticulum (ER), mitochondria, plastids (such as chloroplasts), the vacuole, the Golgi apparatus, protein storage vessicles (PSV) and, in general, membranes. Some signal peptide sequences are conserved, such as the Asn-Pro-Ile-Arg amino acid motif found in the N-terminal propeptide signal that targets proteins to the vacuole (Marty (1999) *The Plant Cell* 11: 587-599). Other signal peptides do not have a consensus sequence *per se*, but are largely composed of hydrophobic amino acids, such as

those signal peptides targeting proteins to the ER (Vitale and Denecke (1999) *The Plant Cell* 11: 615-628). Still others do not appear to contain either a consensus sequence or an identified common secondary sequence, for instance the chloroplast stromal targeting signal peptides (Keegstra and Cline (1999) *The Plant Cell* 11: 557-570). Furthermore, some targeting peptides are bipartite, directing proteins first to an organelle and then to a membrane within the organelle (e.g. within the thylakoid lumen of the chloroplast; see Keegstra and Cline (1999) *The Plant Cell* 11: 557-570). In addition to the diversity in sequence and secondary structure, placement of the signal peptide is also varied. Proteins destined for the vacuole, for example, have targeting signal peptides found at the N-terminus, at the C-terminus and at a surface location in mature, folded proteins. Signal peptides also serve as ligands for some receptors.

These characteristics of signal proteins can be used to more tightly control the phenotypic expression of introduced SDFs. In particular, associating the appropriate signal sequence with a specific SDF can allow sequestering of the protein in specific organelles (plastids, as an example), secretion outside of the cell, targeting interaction with particular receptors, etc. Hence, the inclusion of signal proteins in constructs involving the SDFs of the invention increases the range of manipulation of SDF phenotypic expression. The nucleotide sequence of the signal peptide can be isolated from characterized genes using common molecular biological techniques or can be synthesized in vitro.

In addition, the native signal peptide sequences, both amino acid and nucleotide, described in the Reference, Sequence, Protein Group or Protein Group Matrix tables can be used to modulate polypeptide transport. Further variants of the native signal peptides described in the Reference, Sequence, Protein Group or Protein Group Matrix tables are contemplated. Insertions, deletions, or substitutions can be made. Such variants will retain at least one of the functions of the native signal peptide as well as exhibiting some degree of sequence identity to the native sequence.

Also, fragments of the signal peptides of the invention are useful and can be fused with other signal peptides of interest to modulate transport of a polypeptide.

## V. Transformation Techniques

A wide range of techniques for inserting exogenous polynucleotides are known for a number of host cells, including, without limitation, bacterial, yeast, mammalian, insect and plant cells.

Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, e.g. Weising et al., *Ann. Rev. Genet.* 22:421 (1988); and Christou, *Euphytica*, v. 85, n.1-3:13-27, (1995).

DNA constructs of the invention may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria (McCormac et al., *Mol. Biotechnol.* 8:199 (1997); Hamilton, *Gene* 200:107 (1997); Salomon et al. *EMBO J.* 3:141 (1984); Herrera-Estrella et al. *EMBO J.* 2:987 (1983).

Microinjection techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al. *EMBO J.* 3:2717 (1984). Electroporation techniques are described in Fromm et al. *Proc. Natl Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein et al. *Nature* 327:773 (1987). *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary or co-integrate vectors, are well described in the scientific literature. See, for example Hamilton, *CM.*, *Gene* 200:107 (1997); Müller et al. *Mol. Gen. Genet.* 207:171 (1987); Komari et al. *Plant J.* 10:165 (1996); Venkateswarlu et al. *Biotechnology* 9:1103 (1991) and Gleave, *AP.*, *Plant Mol. Biol.* 20:1203 (1992); Graves and Goldman, *Plant Mol. Biol.* 7:34 (1986) and Gould et al., *Plant Physiology* 95:426 (1991).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant that possesses the transformed genotype and thus the desired phenotype such as seedlessness. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or

herbicide marker which has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al., *Protoplasts Isolation and Culture* in "Handbook of Plant Cell Culture," pp. 124-176, MacMillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1988. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee et al. *Ann. Rev. of Plant Phys.* 38:467 (1987). Regeneration of monocots (rice) is described by Hosoyama et al. (*Biosci. Biotechnol. Biochem.* 58:1500 (1994)) and by Ghosh et al. (*J. Biotechnol.* 32:1 (1994)). The nucleic acids of the invention can be used to confer desired traits on essentially any plant.

Thus, the invention has use over a broad range of plants, including species from the genera *Anacardium*, *Arachis*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Olea*, *Oryza*, *Panicum*, *Pannesetum*, *Persea*, *Phaseolus*, *Pistachia*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Ricinus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Theobromus*, *Trigonella*, *Triticum*, *Vicia*, *Vitis*, *Vigna*, and, *Zea*.

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

The particular sequences of SDFs identified are provided in the attached Reference and Sequence tables.

## DEFINITIONS

Allelic variant: An "allelic variant" is an alternative form of the same SDF, which resides at the same chromosomal locus in the organism. Allelic variations can occur in any portion of the gene sequence, including regulatory regions. Allelic variants can arise by normal genetic variation in a population. Allelic variants can also be produced by genetic engineering methods. An allelic variant can be one that is found in a naturally occurring plant, including a cultivar or ecotype. An allelic variant may or may not give rise to a phenotypic change, and may or may not be expressed. An allele can result in a detectable change in the phenotype of the trait represented by the locus. A phenotypically silent allele can give rise to a product.

Alternatively spliced messages: Within the context of the current invention, "alternatively spliced messages" refers to mature mRNAs originating from a single gene with variations in the number and/or identity of exons, introns and/or intron-exon junctions.

Chimeric: The term "chimeric" is used to describe genes, as defined supra, or constructs wherein at least two of the elements of the gene or construct, such as the promoter and the coding sequence and/or other regulatory sequences and/or filler sequences and/or complements thereof, are heterologous to each other.

Constitutive Promoter: Promoters referred to herein as "constitutive promoters" actively promote transcription under most, but not necessarily all, environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcript initiation region and the 1' or 2' promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes, such as the maize ubiquitin-1 promoter, known to those of skill.

Coordinately Expressed: The term "coordinately expressed," as used in the current invention, refers to genes that are expressed at the same or a similar time and/or stage and/or under the same or similar environmental conditions.

Domain: Domains are fingerprints or signatures that can be used to characterize protein families and/or parts of proteins. Such fingerprints or signatures can comprise conserved (1) primary sequence, (2) secondary structure, and/or (3) three-dimensional conformation. Generally, each domain has been associated with either a family of proteins or motifs. Typically, these families and/or motifs have been correlated with specific *in-vitro* and/or *in-vivo* activities. A domain can be any length, including the entirety of the sequence of a protein. Detailed descriptions of the domains, associated families and motifs, and correlated activities of the polypeptides of the instant invention are described below. Usually, the polypeptides with designated domain(s) can exhibit at least one activity that is exhibited by any polypeptide that comprises the same domain(s).

Endogenous: The term "endogenous," within the context of the current invention refers to any polynucleotide, polypeptide or protein sequence which is a natural part of a cell or organisms regenerated from said cell.

Exogenous: "Exogenous," as referred to within, is any polynucleotide, polypeptide or protein sequence, whether chimeric or not, that is initially or subsequently introduced into the genome of an individual host cell or the organism regenerated from said host cell by any means other than by a sexual cross. Examples of means by which this can be accomplished are described below, and include *Agrobacterium*-mediated transformation (of dicots - *e.g.* Salomon et al. *EMBO J.* 3:141 (1984); Herrera-Estrella et al. *EMBO J.* 2:987 (1983); of monocots, representative papers are those by Escudero et al., *Plant J.* 10:355 (1996), Ishida et al., *Nature Biotechnology* 14:745 (1996), May et al., *Bio/Technology* 13:486 (1995)), biolistic methods (Armaleo et al., *Current Genetics* 17:97 (1990)), electroporation, *in planta* techniques, and the like. Such a plant containing the exogenous nucleic acid is referred to here as a T<sub>0</sub> for the primary transgenic plant and T<sub>1</sub> for the first generation. The term "exogenous" as used herein is also intended to encompass inserting a naturally found element into a non-naturally found location.

Filler sequence: As used herein, "filler sequence" refers to any nucleotide sequence that is inserted into DNA construct to evoke a particular spacing between particular components such as

a promoter and a coding region and may provide an additional attribute such as a restriction enzyme site.

Gene: The term "gene," as used in the context of the current invention, encompasses all regulatory and coding sequence contiguously associated with a single hereditary unit with a genetic function (see SCHEMATIC 1). Genes can include non-coding sequences that modulate the genetic function that include, but are not limited to, those that specify polyadenylation, transcriptional regulation, DNA conformation, chromatin conformation, extent and position of base methylation and binding sites of proteins that control all of these. Genes comprised of "exons" (coding sequences), which may be interrupted by "introns" (non-coding sequences), encode proteins. A gene's genetic function may require only RNA expression or protein production, or may only require binding of proteins and/or nucleic acids without associated expression. In certain cases, genes adjacent to one another may share sequence in such a way that one gene will overlap the other. A gene can be found within the genome of an organism, artificial chromosome, plasmid, vector, etc., or as a separate isolated entity.

Gene Family: "Gene family" is used in the current invention to describe a group of functionally related genes, each of which encodes a separate protein.

Heterologous sequences: "Heterologous sequences" are those that are not operatively linked or are not contiguous to each other in nature. For example, a promoter from corn is considered heterologous to an *Arabidopsis* coding region sequence. Also, a promoter from a gene encoding a growth factor from corn is considered heterologous to a sequence encoding the corn receptor for the growth factor. Regulatory element sequences, such as UTRs or 3' end termination sequences that do not originate in nature from the same gene as the coding sequence originates from, are considered heterologous to said coding sequence. Elements operatively linked in nature and - contiguous to each other are not heterologous to each other. On the other hand, these same elements remain operatively linked but become heterologous if other filler sequence is placed between them. Thus, the promoter and coding sequences of a corn gene expressing an amino acid transporter are not heterologous to each other, but the promoter and coding sequence of a corn gene operatively linked in a novel manner are heterologous.



Homologous gene: In the current invention, "homologous gene" refers to a gene that shares sequence similarity with the gene of interest. This similarity may be in only a fragment of the sequence and often represents a functional domain such as, examples including without limitation a DNA binding domain, a domain with tyrosine kinase activity, or the like. The functional activities of homologous genes are not necessarily the same.

Inducible Promoter: An "inducible promoter" in the context of the current invention refers to a promoter which is regulated under certain conditions, such as light, chemical concentration, protein concentration, conditions in an organism, cell, or organelle, etc. A typical example of an inducible promoter, which can be utilized with the polynucleotides of the present invention, is PARSK1, the promoter from the *Arabidopsis* gene encoding a serine-threonine kinase enzyme, and which promoter is induced by dehydration, abscissic acid and sodium chloride (Wang and Goodman, *Plant J.* 8:37 (1995)) Examples of environmental conditions that may affect transcription by inducible promoters include anaerobic conditions, elevated temperature, or the presence of light.

Intergenic region: "Intergenic region," as used in the current invention, refers to nucleotide sequence occurring in the genome that separates adjacent genes.

Mutant gene: In the current invention, "mutant" refers to a heritable change in DNA sequence at a specific location. Mutants of the current invention may or may not have an associated identifiable function when the mutant gene is transcribed.

Orthologous Gene: In the current invention "orthologous gene" refers to a second gene that encodes a gene product that performs a similar function as the product of a first gene. The orthologous gene may also have a degree of sequence similarity to the first gene. The orthologous gene may encode a polypeptide that exhibits a degree of sequence similarity to a polypeptide corresponding to a first gene. The sequence similarity can be found within a functional domain or along the entire length of the coding sequence of the genes and/or their corresponding polypeptides.

Percentage of sequence identity: "Percentage of sequence identity," as used herein, is determined by comparing two optimally aligned sequences over a comparison window, where the fragment of the polynucleotide or amino acid sequence in the comparison window may comprise additions or deletions (e.g., gaps or overhangs) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Add. APL. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman and Wunsch *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson and Lipman *Proc. Natl. Acad. Sci. (USA)* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, PASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection. Given that two sequences have been identified for comparison, GAP and BESTFIT are preferably employed to determine their optimal alignment. Typically, the default values of 5.00 for gap weight and 0.30 for gap weight length are used. The term "substantial sequence identity" between polynucleotide or polypeptide sequences refers to polynucleotide or polypeptide comprising a sequence that has at least 80% sequence identity, preferably at least 85%, more preferably at least 90% and most preferably at least 95%, even more preferably, at least 96%, 97%, 98% or 99% sequence identity compared to a reference sequence using the programs.

Plant Promoter: A "plant promoter" is a promoter capable of initiating transcription in plant cells and can drive or facilitate transcription of a fragment of the SDF of the instant invention or a coding sequence of the SDF of the instant invention. Such promoters need not be of plant origin. For example, promoters derived from plant viruses, such as the CaMV35S promoter or from *Agrobacterium tumefaciens* such as the T-DNA promoters, can be plant promoters. A typical example of a plant promoter of plant origin is the maize ubiquitin-1 (ubi-1) promoter known to those of skill.

Promoter: The term "promoter," as used herein, refers to a region of sequence determinants located upstream from the start of transcription of a gene and which are involved in recognition and binding of RNA polymerase and other proteins to initiate and modulate transcription. A basal promoter is the minimal sequence necessary for assembly of a transcription complex required for transcription initiation. Basal promoters frequently include a "TATA box" element usually located between 15 and 35 nucleotides upstream from the site of initiation of transcription. Basal promoters also sometimes include a "CCAAT box" element (typically a sequence CCAAT) and/or a GGGCG sequence, usually located between 40 and 200 nucleotides, preferably 60 to 120 nucleotides, upstream from the start site of transcription.

Public sequence: The term "public sequence," as used in the context of the instant application, refers to any sequence that has been deposited in a publicly accessible database. This term encompasses both amino acid and nucleotide sequences. Such sequences are publicly accessible, for example, on the BLAST databases on the NCBI FTP web site (accessible at [ncbi.nlm.gov/blast](http://ncbi.nlm.gov/blast)). The database at the NCBI GTP site utilizes "gi" numbers assigned by NCBI as a unique identifier for each sequence in the databases, thereby providing a non-redundant database for sequence from various databases, including GenBank, EMBL, DDBJ, (DNA Database of Japan) and PDB (Brookhaven Protein Data Bank).

Regulatory Sequence: The term "regulatory sequence," as used in the current invention, refers to any nucleotide sequence that influences transcription or translation initiation and rate, and stability and/or mobility of the transcript or polypeptide product. Regulatory sequences include, but are not limited to, promoters, promoter control elements, protein binding sequences, 5' and 3' UTRs, transcriptional start site, termination sequence, polyadenylation sequence, introns, certain sequences within a coding sequence, etc.

Related Sequences: "Related sequences" refer to either a polypeptide or a nucleotide sequence that exhibits some degree of sequence similarity with a sequence described by The Reference tables and The Sequence tables.

Scaffold Attachment Region (SAR): As used herein, "scaffold attachment region" is a DNA sequence that anchors chromatin to the nuclear matrix or scaffold to generate loop domains that can have either a transcriptionally active or inactive structure (Spiker and Thompson (1996) *Plant Physiol.* 110: 15-21).

Sequence-determined DNA fragments (SDFs): "Sequence-determined DNA fragments" as used in the current invention are isolated sequences of genes, fragments of genes, intergenic regions or contiguous DNA from plant genomic DNA or cDNA or RNA the sequence of which has been determined.

Signal Peptide: A "signal peptide" as used in the current invention is an amino acid sequence that targets the protein for secretion, for transport to an intracellular compartment or organelle or for incorporation into a membrane. Signal peptides are indicated in the tables and a more detailed description located below.

Specific Promoter: In the context of the current invention, "specific promoters" refers to a subset of inducible promoters that have a high preference for being induced in a specific tissue or cell and/or at a specific time during development of an organism. By "high preference" is meant at least 3-fold, preferably 5-fold, more preferably at least 10-fold still more preferably at least 20-fold, 50-fold or 100-fold increase in transcription in the desired tissue over the transcription in any other tissue. Typical examples of temporal and/or tissue specific promoters of plant origin that can be used with the polynucleotides of the present invention, are: PTA29, a promoter which is capable of driving gene transcription specifically in tapetum and only during anther development (Koltonow et al., *Plant Cell* 2:1201 (1990); RCc2 and RCc3, promoters that direct root-specific gene transcription in rice (Xu et al., *Plant Mol. Biol.* 27:237 (1995); TobRB27, a root-specific promoter from tobacco (Yamamoto et al., *Plant Cell* 3:371 (1991)). Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only in certain tissues or organs, such as root, ovule, fruit, seeds, or flowers. Other suitable promoters include those from genes encoding storage proteins or the lipid body membrane protein, oleosin. A few root-specific promoters are noted above.

Stringency: "Stringency" as used herein is a function of probe length, probe composition (G + C content), and salt concentration, organic solvent concentration, and temperature of hybridization or wash conditions. Stringency is typically compared by the parameter  $T_m$ , which is the temperature at which 50% of the complementary molecules in the hybridization are hybridized, in terms of a temperature differential from  $T_m$ . High stringency conditions are those providing a condition of  $T_m - 5^\circ\text{C}$  to  $T_m - 10^\circ\text{C}$ . Medium or moderate stringency conditions are those providing  $T_m - 20^\circ\text{C}$  to  $T_m - 29^\circ\text{C}$ . Low stringency conditions are those providing a condition of  $T_m - 40^\circ\text{C}$  to  $T_m - 48^\circ\text{C}$ . The relationship of hybridization conditions to  $T_m$  (in  $^\circ\text{C}$ ) is expressed in the mathematical equation

$$T_m = 81.5 - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\%G+C) - (600/N) \quad (1)$$

where N is the length of the probe. This equation works well for probes 14 to 70 nucleotides in length that are identical to the target sequence. The equation below for  $T_m$  of DNA-DNA hybrids is useful for probes in the range of 50 to greater than 500 nucleotides, and for conditions that include an organic solvent (formamide).

$$T_m = 81.5 + 16.6 \log \{ [\text{Na}^+]/(1 + 0.7[\text{Na}^+]) \} + 0.41(\%G+C) - 500/L - 0.63(\%\text{formamide}) \quad (2)$$

where L is the length of the probe in the hybrid. (P. Tijessen, "Hybridization with Nucleic Acid Probes" in Laboratory Techniques in Biochemistry and Molecular Biology, P.C. van der Vliet, ed., c. 1993 by Elsevier, Amsterdam.) The  $T_m$  of equation (2) is affected by the nature of the hybrid; for DNA-RNA hybrids  $T_m$  is  $10-15^\circ\text{C}$  higher than calculated, for RNA-RNA hybrids  $T_m$  is  $20-25^\circ\text{C}$  higher. Because the  $T_m$  decreases about  $1^\circ\text{C}$  for each 1% decrease in homology when a long probe is used (Bonner et al., *J. Mol. Biol.* 81:123 (1973)), stringency conditions can be adjusted to favor detection of identical genes or related family members.

Equation (2) is derived assuming equilibrium and therefore, hybridizations according to the present invention are most preferably performed under conditions of probe excess and for sufficient time to achieve equilibrium. The time required to reach equilibrium can be shortened by inclusion of a hybridization accelerator such as dextran sulfate or another high volume polymer in the hybridization buffer.

Stringency can be controlled during the hybridization reaction or after hybridization has occurred by altering the salt and temperature conditions of the wash solutions used. The formulas shown above are equally valid when used to compute the stringency of a wash solution. Preferred wash solution stringencies lie within the ranges stated above; high stringency is 5-8°C below  $T_m$ , medium or moderate stringency is 26-29°C below  $T_m$  and low stringency is 45-48°C below  $T_m$ .

Substantially free of: A composition containing A is "substantially free of" B when at least 85% by weight of the total A+B in the composition is A. Preferably, A comprises at least about 90% by weight of the total of A+B in the composition, more preferably at least about 95% or even 99% by weight. For example, a plant gene or DNA sequence can be considered substantially free of other plant genes or DNA sequences.

Translational start site: In the context of the current invention, a "translational start site" is usually an ATG in the cDNA transcript, more usually the first ATG. A single cDNA, however, may have multiple translational start sites.

Transcription start site: "Transcription start site" is used in the current invention to describe the point at which transcription is initiated. This point is typically located about 25 nucleotides downstream from a TFIID binding site, such as a TATA box. Transcription can initiate at one or more sites within the gene, and a single gene may have multiple transcriptional start sites, some of which may be specific for transcription in a particular cell-type or tissue.

Untranslated region (UTR): A "UTR" is any contiguous series of nucleotide bases that is transcribed, but is not translated. These untranslated regions may be associated with particular functions such as increasing mRNA message stability. Examples of UTRs include, but are not limited to polyadenylation signals, terminations sequences, sequences located between the transcriptional start site and the first exon (5' UTR) and sequences located between the last exon and the end of the mRNA (3' UTR).

Variant: The term "variant" is used herein to denote a polypeptide or protein or polynucleotide molecule that differs from others of its kind in some way. For example, polypeptide and protein variants can consist of changes in amino acid sequence and/or charge and/or post-translational modifications (such as glycosylation, etc).

## EXAMPLES

The invention is illustrated by way of the following examples. The invention is not limited by these examples as the scope of the invention is defined solely by the claims following.

### EXAMPLE 1: cDNA PREPARATION

5 A number of the nucleotide sequences disclosed in the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables, herein as representative of the SDFs of the invention can be obtained by sequencing genomic DNA (gDNA) and/or cDNA from corn plants grown from HYBRID SEED # 35A19, purchased from Pioneer Hi-Bred International, Inc., Supply Management, P.O. Box 256, Johnston, Iowa 50131-0256.

A number of the nucleotide sequences disclosed in the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables, herein as representative of the SDFs of the invention can also be obtained by sequencing genomic DNA from *Arabidopsis thaliana*, Wassilewskija ecotype or by sequencing cDNA obtained from mRNA from such plants as described below. This is a true breeding strain. Seeds of the plant are available from the Arabidopsis Biological Resource Center at the Ohio State University, under the accession number CS2360. Seeds of this plant were deposited under the terms and conditions of the Budapest Treaty at the American Type Culture Collection, Manassas, VA on August 31, 1999, and were assigned ATCC No. PTA-595.

10 Other methods for cloning full-length cDNA are described, for example, by Seki et al., *Plant Journal* 15:707-720 (1998) "High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated Cap trapper"; Maruyama et al., *Gene* 138:171 (1994) "Oligo-capping a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides"; and WO 96/34981.

25 Tissues were, or each organ was, individually pulverized and frozen in liquid nitrogen. Next, the samples were homogenized in the presence of detergents and then centrifuged. The debris and nuclei were removed from the sample and more detergents were added to the sample. The sample was centrifuged and the debris was removed. Then the sample was applied to a 2M sucrose cushion to isolate polysomes. The RNA was isolated by treatment with detergents and proteinase K followed by ethanol precipitation and centrifugation. The polysomal RNA from the

30



different tissues was pooled according to the following mass ratios: 15/15/1 for male inflorescences, female inflorescences and root, respectively. The pooled material was then used for cDNA synthesis by the methods described below.

Starting material for cDNA synthesis for the exemplary corn cDNA clones with sequences presented in the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables was poly(A)-containing polysomal mRNAs from inflorescences and root tissues of corn plants grown from HYBRID SEED # 35A19. Male inflorescences and female (pre-and post-fertilization) inflorescences were isolated at various stages of development. Selection for poly(A) containing polysomal RNA was done using oligo d(T) cellulose columns, as described by Cox and Goldberg, "Plant Molecular Biology: A Practical Approach", pp. 1-35, Shaw ed., c. 1988 by IRL, Oxford. The quality and the integrity of the polyA+ RNAs were evaluated.

Starting material for cDNA synthesis for the exemplary *Arabidopsis* cDNA clones with sequences presented in the Reference and Sequence tables or polynucleotides encoding polypeptides of the Protein Group or Protein Group Matrix tables was polysomal RNA isolated from the top-most inflorescence tissues of *Arabidopsis thaliana* Wassilewskija (Ws.) and from roots of *Arabidopsis thaliana* Landsberg erecta (L. er.), also obtained from the Arabidopsis Biological Resource Center. Nine parts inflorescence to every part root was used, as measured by wet mass. Tissue was pulverized and exposed to liquid nitrogen. Next, the sample was homogenized in the presence of detergents and then centrifuged. The debris and nuclei were removed from the sample and more detergents were added to the sample. The sample was centrifuged and the debris was removed and the sample was applied to a 2M sucrose cushion to isolate polysomal RNA. Cox et al., "Plant Molecular Biology: A Practical Approach", pp. 1-35, Shaw ed., c. 1988 by IRL, Oxford. The polysomal RNA was used for cDNA synthesis by the methods described below. Polysomal mRNA was then isolated as described above for corn cDNA. The quality of the RNA was assessed electrophoretically.

Following preparation of the mRNAs from various tissues as described above, selection of mRNA with intact 5' ends and specific attachment of an oligonucleotide tag to the 5' end of such mRNA was performed using either a chemical or enzymatic approach. Both techniques take advantage of the presence of the "cap" structure, which characterizes the 5' end of most intact mRNAs and which comprises a guanosine generally methylated once, at the 7 position.

The chemical modification approach involves the optional elimination of the 2', 3'-cis diol of the 3' terminal ribose, the oxidation of the 2', 3'-cis diol of the ribose linked to the cap of the 5' ends of the mRNAs into a dialdehyde, and the coupling of the such obtained dialdehyde to a derivatized oligonucleotide tag. Further detail regarding the chemical approaches for obtaining mRNAs having intact 5' ends are disclosed in International Application No. WO96/34981 published November 7, 1996.

The enzymatic approach for ligating the oligonucleotide tag to the intact 5' ends of mRNAs involves the removal of the phosphate groups present on the 5' ends of uncapped incomplete mRNAs, the subsequent decapping of mRNAs having intact 5' ends and the ligation of the phosphate present at the 5' end of the decapped mRNA to an oligonucleotide tag. Further detail regarding the enzymatic approaches for obtaining mRNAs having intact 5' ends are disclosed in Dumas Milne Edwards J.B. (Doctoral Thesis of Paris VI University, Le clonage des ADNc complets: difficultés et perspectives nouvelles. Apports pour l'étude de la régulation de l'expression de la tryptophane hydroxylase de rat, 20 Dec. 1993), EP0 625572 and Kato *et al.*, *Gene* 150:243-250 (1994).

In both the chemical and the enzymatic approach, the oligonucleotide tag has a restriction enzyme site (e.g. an EcoRI site) therein to facilitate later cloning procedures. Following attachment of the oligonucleotide tag to the mRNA, the integrity of the mRNA is examined by performing a Northern blot using a probe complementary to the oligonucleotide tag.

For the mRNAs joined to oligonucleotide tags using either the chemical or the enzymatic method, first strand cDNA synthesis is performed using an oligo-dT primer with reverse transcriptase. This oligo-dT primer can contain an internal tag of at least 4 nucleotides, which can be different from one mRNA preparation to another. Methylated dCTP is used for cDNA first strand synthesis to protect the internal EcoRI sites from digestion during subsequent steps. The first strand cDNA is precipitated using isopropanol after removal of RNA by alkaline hydrolysis to eliminate residual primers.

Second strand cDNA synthesis is conducted using a DNA polymerase, such as Klenow fragment and a primer corresponding to the 5' end of the ligated oligonucleotide. The primer is typically 20-25 bases in length. Methylated dCTP is used for second strand synthesis in order to protect internal EcoRI sites in the cDNA from digestion during the cloning process.

Following second strand synthesis, the full-length cDNAs are cloned into a phagemid vector, such as pBlueScript™ (Stratagene). The ends of the full-length cDNAs are blunted with T4 DNA polymerase (Biolabs) and the cDNA is digested with EcoRI. Since methylated dCTP is used during cDNA synthesis, the EcoRI site present in the tag is the only hemi-methylated site; hence the only site susceptible to EcoRI digestion. In some instances, to facilitate subcloning, an Hind III adapter is added to the 3' end of full-length cDNAs.

The full-length cDNAs are then size fractionated using either exclusion chromatography (AcA, Biosepra) or electrophoretic separation which yields 3 to 6 different fractions. The full-length cDNAs are then directionally cloned either into pBlueScript™ using either the EcoRI and SmaI restriction sites or, when the Hind III adapter is present in the full-length cDNAs, the EcoRI and Hind III restriction sites. The ligation mixture is transformed, preferably by electroporation, into bacteria, which are then propagated under appropriate antibiotic selection.

Clones containing the oligonucleotide tag attached to full-length cDNAs are selected as follows.

The plasmid cDNA libraries made as described above are purified (e.g. by a column available from Qiagen). A positive selection of the tagged clones is performed as follows. Briefly, in this selection procedure, the plasmid DNA is converted to single stranded DNA using phage F1 gene II endonuclease in combination with an exonuclease (Chang et al., *Gene* 127:95 (1993)) such as exonuclease III or T7 gene 6 exonuclease. The resulting single stranded DNA is then purified using paramagnetic beads as described by Fry et al., *Biotechniques* 13: 124 (1992). Here the single stranded DNA is hybridized with a biotinylated oligonucleotide having a sequence corresponding to the 3' end of the oligonucleotide tag. Preferably, the primer has a length of 20-25 bases. Clones including a sequence complementary to the biotinylated oligonucleotide are selected by incubation with streptavidin coated magnetic beads followed by magnetic capture. After capture of the positive clones, the plasmid DNA is released from the magnetic beads and converted into double stranded DNA using a DNA polymerase such as ThermoSequenase™ (obtained from Amersham Pharmacia Biotech). Alternatively, protocols such as the Gene Trapper™ kit (Gibco BRL) can be used. The double stranded DNA is then transformed, preferably by electroporation, into bacteria. The percentage of positive clones having the 5' tag oligonucleotide is typically estimated to be between 90 and 98% from dot blot analysis.

Following transformation, the libraries are ordered in microtiter plates and sequenced. The *Arabidopsis* library was deposited at the American Type Culture Collection on January 7, 2000 as "*E-coli* liba 010600" under the accession number PTA-1161.

G. EXAMPLE 2: Southern hybridizations

The SDFs of the invention can be used in Southern hybridizations as described above. The following describes extraction of DNA from nuclei of plant cells, digestion of the nuclear DNA and separation by length, transfer of the separated fragments to membranes, preparation of probes for hybridization, hybridization and detection of the hybridized probe.

The procedures described herein can be used to isolate related polynucleotides or for diagnostic purposes. Moderate stringency hybridization conditions, as defined above, are described in the present example. These conditions result in detection of hybridization between sequences having at least 70% sequence identity. As described above, the hybridization and wash conditions can be changed to reflect the desired percentatge of sequence identity between probe and target sequences that can be detected.

In the following procedure, a probe for hybridization is produced from two PCR reactions using two primers from genomic sequence of *Arabidopsis thaliana*. As described above, the particular template for generating the probe can be any desired template.

The first PCR product is assessed to validate the size of the primer to assure it is of the expected size. Then the product of the first PCR is used as a template, with the same pair of primers used in the first PCR, in a second PCR that produces a labeled product used as the probe.

Fragments detected by hybridization, or other bands of interest, can be isolated from gels used to separate genomic DNA fragments by known methods for further purification and/or characterization.

Buffers for nuclear DNA extraction

1. 10X HB

	1000 ml	
40 mM spermidine	10.2 g	Spermine (Sigma S-2876) and spermidine (Sigma S-2501)

10 mM spermine	3.5 g	Stabilize chromatin and the nuclear membrane
0.1 M EDTA (disodium)	37.2 g	EDTA inhibits nuclease
0.1 M Tris	12.1 g	Buffer
0.8 M KCl	59.6 g	Adjusts ionic strength for stability of nuclei

Adjust pH to 9.5 with 10 N NaOH. It appears that there is a nuclease present in leaves.  
Use of pH 9.5 appears to inactivate this nuclease.

2. 2 M sucrose (684 g per 1000 ml)  
Heat about half the final volume of water to about 50°C. Add the sucrose slowly then bring the mixture to close to final volume; stir constantly until it has dissolved. Bring the solution to volume.
3. Sarkosyl solution (lyses nuclear membranes)

	<u>1000 ml</u>
N-lauroyl sarcosine (Sarkosyl)	20.0 g
0.1 M Tris	12.1 g
0.04 M EDTA (Disodium)	14.9 g

Adjust the pH to 9.5 after all the components are dissolved and bring up to the proper volume.

4. 20% Triton X-100  
80 ml Triton X-100  
320 ml 1xHB (w/o  $\beta$ -ME and PMSF)

Prepare in advance; Triton takes some time to dissolve

A. Procedure

1. Prepare 1X "H" buffer (keep ice-cold during use)

	<u>1000 ml</u>
10X HB	100 ml
2 M sucrose	250 ml a non-ionic osmoticum
Water	634 ml

Added just before use:

100 mM PMSF*	10 ml a protease inhibitor; protects nuclear membrane proteins
$\beta$ -mercaptoethanol	1 ml inactivates nuclease by reducing disulfide bonds

\*100 mM PMSF  
(phenyl methyl sulfonyl fluoride, Sigma P-7626)  
(add 0.0875 g to 5 ml 100% ethanol)

2. Homogenize the tissue in a blender (use 300-400 ml of 1xHB per blender). Be sure that you use 5-10 ml of HB buffer per gram of tissue. Blenders generate heat so be sure to keep the homogenate cold. It is necessary to put the blenders in ice periodically.
3. Add the 20% Triton X-100 (25 ml per liter of homogenate) and gently stir on ice for 20 min. This lyses plastid, but not nuclear, membranes.
4. Filter the tissue suspension through several nylon filters into an ice-cold beaker. The first filtration is through a 250-micron membrane; the second is through an 85-micron membrane; the third is through a 50-micron membrane; and the fourth is through a 20-

micron membrane. Use a large funnel to hold the filters. Filtration can be sped up by gently squeezing the liquid through the filters.

5. Centrifuge the filtrate at 1200 x g for 20 min. at 4°C to pellet the nuclei.
6. Discard the dark green supernatant. The pellet will have several layers to it. One is starch; it is white and gritty. The nuclei are gray and soft. In the early steps, there may be a dark green and somewhat viscous layer of chloroplasts.

Wash the pellets in about 25 ml cold H buffer (with Triton X-100) and resuspend by swirling gently and pipetting. After the pellets are resuspended.

Pellet the nuclei again at 1200 - 1300 x g. Discard the supernatant.

Repeat the wash 3-4 times until the supernatant has changed from a dark green to a pale green. This usually happens after 3 or 4 resuspensions. At this point, the pellet is typically grayish white and very slippery. The Triton X-100 in these repeated steps helps to destroy the chloroplasts and mitochondria that contaminate the prep.

Resuspend the nuclei for a final time in a total of 15 ml of H buffer and transfer the suspension to a sterile 125 ml Erlenmeyer flask.

7. Add 15 ml, dropwise, cold 2% Sarkosyl, 0.1 M Tris, 0.04 M EDTA solution (pH 9.5) while swirling gently. This lyses the nuclei. The solution will become very viscous.
8. Add 30 grams of CsCl and gently swirl at room temperature until the CsCl is in solution. The mixture will be gray, white and viscous.
9. Centrifuge the solution at 11,400 x g at 4°C for at least 30 min. The longer this spin is, the firmer the protein pellicle.

10. The result is typically a clear green supernatant over a white pellet, and (perhaps) under a protein pellicle. Carefully remove the solution under the protein pellicle and above the pellet. Determine the density of the solution by weighing 1 ml of solution and add CsCl if necessary to bring to 1.57 g/ml. The solution contains dissolved solids (sucrose etc) and the refractive index alone will not be an accurate guide to CsCl concentration.
11. Add 20  $\mu$ l of 10 mg/ml EtBr per ml of solution.
12. Centrifuge at 184,000 x g for 16 to 20 hours in a fixed-angle rotor.
13. Remove the dark red supernatant that is at the top of the tube with a plastic transfer pipette and discard. Carefully remove the DNA band with another transfer pipette. The DNA band is usually visible in room light; otherwise, use a long wave UV light to locate the band.
14. Extract the ethidium bromide with isopropanol saturated with water and salt. Once the solution is clear, extract at least two more times to ensure that all of the EtBr is gone. Be very gentle, as it is very easy to shear the DNA at this step. This extraction may take a while because the DNA solution tends to be very viscous. If the solution is too viscous, dilute it with TE.
15. Dialyze the DNA for at least two days against several changes (at least three times) of TE (10 mM Tris, 1mM EDTA, pH 8) to remove the cesium chloride.
16. Remove the dialyzed DNA from the tubing. If the dialyzed DNA solution contains a lot of debris, centrifuge the DNA solution at least at 2500 x g for 10 min. and carefully transfer the clear supernatant to a new tube. Read the A260 concentration of the DNA.
17. Assess the quality of the DNA by agarose gel electrophoresis (1% agarose gel) of the DNA. Load 50 ng and 100 ng (based on the OD reading) and compare it with known and



good quality DNA. Undigested lambda DNA and a lambda-HindIII-digested DNA are good molecular weight makers.

#### Protocol for Digestion of Genomic DNA

##### Protocol:

1. The relative amounts of DNA for different crop plants that provide approximately a balanced number of genome equivalent is given in Table 3. Note that due to the size of the wheat genome, wheat DNA will be underrepresented. Lambda DNA provides a useful control for complete digestion.
2. Precipitate the DNA by adding 3 volumes of 100% ethanol. Incubate at  $-20^{\circ}\text{C}$  for at least two hours. Yeast DNA can be purchased and made up at the necessary concentration, therefore no precipitation is necessary for yeast DNA.
3. Centrifuge the solution at  $11,400 \times g$  for 20 min. Decant the ethanol carefully (be careful not to disturb the pellet). Be sure that the residual ethanol is completely removed either by vacuum desiccation or by carefully wiping the sides of the tubes with a clean tissue.
4. Resuspend the pellet in an appropriate volume of water. Be sure the pellet is fully resuspended before proceeding to the next step. This may take about 30 min.
5. Add the appropriate volume of 10X reaction buffer provided by the manufacturer of the restriction enzyme to the resuspended DNA followed by the appropriate volume of enzymes. Be sure to mix it properly by slowly swirling the tubes.
6. Set-up the lambda digestion-control for each DNA that you are digesting.
7. Incubate both the experimental and lambda digests overnight at  $37^{\circ}\text{C}$ . Spin down condensation in a microfuge before proceeding.

8. After digestion, add 2  $\mu$ l of loading dye (typically 0.25% bromophenol blue, 0.25% xylene cyanol in 15% Ficoll or 30% glycerol) to the lambda-control digests and load in 1% TPE-agarose gel (TPE is 90 mM Tris-phosphate, 2 mM EDTA, pH 8). If the lambda DNA in the lambda control digests are completely digested, proceed with the precipitation of the genomic DNA in the digests.
9. Precipitate the digested DNA by adding 3 volumes of 100% ethanol and incubating in  $-20^{\circ}\text{C}$  for at least 2 hours (preferably overnight).

EXCEPTION: *Arabidopsis* and yeast DNA are digested in an appropriate volume; they don't have to be precipitated.

10. Resuspend the DNA in an appropriate volume of TE (e.g., 22  $\mu$ l x 50 blots = 1100  $\mu$ l) and an appropriate volume of 10X loading dye (e.g., 2.4  $\mu$ l x 50 blots = 120  $\mu$ l). Be careful in pipetting the loading dye - it is viscous. Be sure you are pipetting the correct volume.

Table 3

Some guide points in digesting genomic DNA.

Species	Genome Size	Size Relative to Arabidopsis	Genome Equivalent to 2 $\mu$ g Arabidopsis DNA	Amount of DNA per blot
Arabidopsis	120 Mb	1X	1X	2 $\mu$ g
Brassica	1,100 Mb	9.2X	0.54X	10 $\mu$ g
Corn	2,800 Mb	23.3X	0.43X	20 $\mu$ g

Cotton	2,300 Mb	19.2X	0.52X	20 µg
Oat	11,300 Mb	94X	0.11X	20 µg
Rice	400 Mb	3.3X	0.75X	5 µg
Soybean	1,100 Mb	9.2X	0.54X	10 µg
Sugarbeet	758 Mb	6.3X	0.8X	10 µg
Sweetclover	1,100 Mb	9.2X	0.54X	10 µg
Wheat	16,000 Mb	133X	0.08X	20 µg
Yeast	15 Mb	0.12X	1X	0.25 µg

#### Protocol for Southern Blot Analysis

The digested DNA samples are electrophoresed in 1% agarose gels in 1x TPE buffer. Low voltage; overnight separations are preferred. The gels are stained with EtBr and photographed.

1. For blotting the gels, first incubate the gel in 0.25 N HCl (with gentle shaking) for about 15 min.
2. Then briefly rinse with water. The DNA is denatured by 2 incubations. Incubate (with shaking) in 0.5 M NaOH in 1.5 M NaCl for 15 min.
3. The gel is then briefly rinsed in water and neutralized by incubating twice (with shaking) in 1.5 M Tris pH 7.5 in 1.5 M NaCl for 15 min.
4. A nylon membrane is prepared by soaking it in water for at least 5 min, then in 6X SSC for at least 15 min. before use. (20x SSC is 175.3 g NaCl, 88.2 g sodium citrate per liter, adjusted to pH 7.0.)
5. The nylon membrane is placed on top of the gel and all bubbles in between are removed. The DNA is blotted from the gel to the membrane using an absorbent medium, such as paper toweling and 6x SCC buffer. After the transfer, the membrane may be lightly brushed with a gloved hand to remove any agarose sticking to the surface.

6. The DNA is then fixed to the membrane by UV crosslinking and baking at 80°C. The membrane is stored at 4°C until use.

B. Protocol for PCR Amplification of Genomic Fragments in Arabidopsis

Amplification procedures:

1. Mix the following in a 0.20 ml PCR tube or 96-well PCR plate:

Volume	Stock	Final Amount or Conc.
0.5 µl	~ 10 ng/µl genomic DNA <sup>1</sup>	5 ng
2.5 µl	10X PCR buffer	20 mM Tris, 50 mM KCl
0.75 µl	50 mM MgCl <sub>2</sub>	1.5 mM
1 µl	10 pmol/µl Primer 1 (Forward)	10 pmol
1 µl	10 pmol/µl Primer 2 (Reverse)	10 pmol
0.5 µl	5 mM dNTPs	0.1 mM
0.1 µl	5 units/µl Platinum Taq™ (Life Technologies, Gaithersburg, MD) DNA Polymerase	1 units
(to 25 µl)	Water	

<sup>1</sup> Arabidopsis DNA is used in the present experiment, but the procedure is a general one.

2. The template DNA is amplified using a Perkin Elmer 9700 PCR machine:

1) 94°C for 10 min. followed by

2) 5 cycles:	3) 5 cycles:	4) 25 cycles:
94 °C - 30 sec	94 °C - 30 sec	94 °C - 30 sec
62 °C - 30 sec	58 °C - 30 sec	53 °C - 30 sec
72 °C - 3 min	72 °C - 3 min	72 °C - 3 min

5) 72°C for 7 min. Then the reactions are stopped by chilling to 4°C.

The procedure can be adapted to a multi-well format if necessary.

#### Quantification and Dilution of PCR Products:

1. The product of the PCR is analyzed by electrophoresis in a 1% agarose gel. A linearized plasmid DNA can be used as a quantification standard (usually at 50, 100, 200, and 400 ng). These will be used as references to approximate the amount of PCR products. HindIII-digested Lambda DNA is useful as a molecular weight marker. The gel can be run fairly quickly; e.g., at 100 volts. The standard gel is examined to determine that the size of the PCR products is consistent with the expected size and if there are significant extra bands or smeary products in the PCR reactions.

2. The amounts of PCR products can be estimated on the basis of the plasmid standard.

3. For the small number of reactions that produce extraneous bands, a small amount of DNA from bands with the correct size can be isolated by dipping a sterile 10-μl tip into the band while viewing through a UV Transilluminator. The small amount of agarose gel (with the DNA fragment) is used in the labeling reaction.

C. Protocol for PCR-DIG-Labeling of DNA

Solutions:

Reagents in PCR reactions (diluted PCR products, 10X PCR Buffer, 50 mM MgCl<sub>2</sub>, 5 U/μl Platinum Taq Polymerase, and the primers)

5 10X dNTP + DIG-11-dUTP [1:5]: (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.65 mM dTTP, 0.35 mM DIG-11-dUTP)

10X dNTP + DIG-11-dUTP [1:10]: (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.81 mM dTTP, 0.19 mM DIG-11-dUTP)

10X dNTP + DIG-11-dUTP [1:15]: (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.875 mM dTTP, 0.125 mM DIG-11-dUTP)

TE buffer (10 mM Tris, 1 mM EDTA, pH 8)

Maleate buffer: In 700 ml of deionized distilled water, dissolve 11.61 g maleic acid and 8.77 g NaCl. Add NaOH to adjust the pH to 7.5. Bring the volume to 1 L. Stir for 15 min. and sterilize.

15 10% blocking solution: In 80 ml deionized distilled water, dissolve 1.16g maleic acid. Next, add NaOH to adjust the pH to 7.5. Add 10 g of the blocking reagent powder (Boehringer Mannheim, Indianapolis, IN, Cat. no. 1096176). Heat to 60°C while stirring to dissolve the powder. Adjust the volume to 100 ml with water. Stir and sterilize.

1% blocking solution: Dilute the 10% stock to 1% using the maleate buffer.

20 Buffer 3 (100 mM Tris, 100 mM NaCl, 50 mM MgCl<sub>2</sub>, pH9.5). Prepared from autoclaved solutions of 1M Tris pH 9.5, 5 M NaCl, and 1 M MgCl<sub>2</sub> in autoclaved distilled water.

Procedure:

1. PCR reactions are performed in 25  $\mu$ l volumes containing:

PCR buffer	1X
MgCl <sub>2</sub>	1.5 mM
10X dNTP + DIG-11-dUTP	1X (please see the note below)
Platinum Taq™ Polymerase	1 unit
10 pg probe DNA	
10 pmol primer 1	

Note: Use for:

10X dNTP + DIG-11-dUTP (1:5)	< 1 kb
10X dNTP + DIG-11-dUTP (1:10)	1 kb to 1.8 kb
10X dNTP + DIG-11-dUTP (1:15)	> 1.8 kb

2. The PCR reaction uses the following amplification cycles:

- 1) 94°C for 10 min.

2) 5 cycles:	3) 5 cycles:	4) 25 cycles:
95°C - 30 sec 61°C - 1 min 73°C - 5 min	95°C - 30 sec 59°C - 1 min 75°C - 5 min	95°C - 30 sec 51°C - 1 min 73°C - 5 min

- 5) 72°C for 8 min. The reactions are terminated by chilling to 4°C (hold).

3. The products are analyzed by electrophoresis- in a 1% agarose gel, comparing to an aliquot of the unlabelled probe starting material.

4. The amount of DIG-labeled probe is determined as follows:

Make serial dilutions of the diluted control DNA in dilution buffer (TE: 10 mM Tris and 1 mM EDTA, pH 8) as shown in the following table:

DIG-labeled control DNA starting conc.	Stepwise Dilution	Final Conc. (Dilution Name)
5 ng/ $\mu$ l	1 $\mu$ l in 49 $\mu$ l TE	100 pg/ $\mu$ l (A)
100 pg/ $\mu$ l (A)	25 $\mu$ l in 25 $\mu$ l TE	50 pg/ $\mu$ l (B)
50 pg/ $\mu$ l (B)	25 $\mu$ l in 25 $\mu$ l TE	25 pg/ $\mu$ l (C)
25 pg/ $\mu$ l (C)	20 $\mu$ l in 30 $\mu$ l TE	10 pg/ $\mu$ l (D)

- a. Serial dilutions of a DIG-labeled standard DNA ranging from 100 pg to 10 pg are spotted onto a positively charged nylon membrane, marking the membrane lightly with a pencil to identify each dilution.
- b. Serial dilutions (e.g., 1:50, 1:2500, 1:10,000) of the newly labeled DNA probe are spotted.
- c. The membrane is fixed by UV crosslinking.
- d. The membrane is wetted with a small amount of maleate buffer and then incubated in 1% blocking solution for 15 min at room temp.
- e. The labeled DNA is then detected using alkaline phosphatase conjugated anti-DIG antibody (Boehringer Mannheim, Indianapolis, IN, cat. no. 1093274) and an NBT substrate according to the manufacture's instruction.



- f. Spot intensities of the control and experimental dilutions are then compared to estimate the concentration of the PCR-DIG-labeled probe.

D. Prehybridization and Hybridization of Southern Blots

Solutions:

5            100% Formamide            purchased from Gibco

20X SSC            (1X = 0.15 M NaCl, 0.015 M Na<sub>3</sub>citrate)

per L:            175 g NaCl

87.5 g Na<sub>3</sub>citrate·2H<sub>2</sub>O

20% Sarkosyl (N-lauroyl-sarcosine)

20% SDS (sodium dodecyl sulphate)

10% Blocking Reagent: In 80 ml deionized distilled water, dissolve 1.16 g maleic acid.  
Next, add NaOH to adjust the pH to 7.5. Add 10 g of the blocking reagent powder.  
Heat to 60°C while stirring to dissolve the powder. Adjust the volume to 100 ml  
with water. Stir and sterilize.

15 Prehybridization Mix:

Final Concentration	Components	Volume (per 100 ml)	Stock
50%	Formamide	50 ml	100%
5X	SSC	25 ml	20X
0.1%	Sarkosyl	0.5 ml	20%
0.02%	SDS	0.1 ml	20%
2%	Blocking Reagent	20 ml	10%
	Water	4.4 ml	

General Procedures:

1. Place the blot in a heat-sealable plastic bag and add an appropriate volume of prehybridization solution (30 ml/100cm<sup>2</sup>) at room temperature. Seal the bag with a heat sealer, avoiding bubbles as much as possible. Lay down the bags in a large plastic tray (one tray can accommodate at least 4–5 bags). Ensure that the bags are lying flat in the tray so that the prehybridization solution is evenly distributed throughout the bag. Incubate the blot for at least 2 hours with gentle agitation using a waver shaker.
2. Denature DIG-labeled DNA probe by incubating for 10 min. at 98°C using the PCR machine and immediately cool it to 4°C.
3. Add probe to prehybridization solution (25 ng/ml; 30 ml = 750 ng total probe) and mix well but avoid foaming. Bubbles may lead to background.
4. Pour off the prehybridization solution from the hybridization bags and add new prehybridization and probe solution mixture to the bags containing the membrane.
5. Incubate with gentle agitation for at least 16 hours.
6. Proceed to medium stringency post-hybridization wash:

Three times for 20 min. each with gentle agitation using 1X SSC, 1% SDS at 60°C.

All wash solutions must be prewarmed to 60°C. Use about 100 ml of wash solution per membrane.

To avoid background keep the membranes fully submerged to avoid drying in spots; agitate sufficiently to avoid having membranes stick to one another.

7. After the wash, proceed to immunological detection and CSPD development.

E. Procedure for Immunological Detection with CSPD

Solutions:

Buffer 1:	Maleic acid buffer (0.1 M maleic acid, 0.15 M NaCl; adjusted to pH 7.5 with NaOH)
5 Washing buffer:	Maleic acid buffer with 0.3% (v/v) Tween 20.
Blocking stock solution	10% blocking reagent in buffer 1. Dissolve (10X concentration): blocking reagent powder (Boehringer Mannheim, Indianapolis, IN, cat. no. 1096176) by constantly stirring on a 65°C heating block or heat in a microwave, autoclave and store at 4°C.
Buffer 2 (1X blocking solution):	Dilute the stock solution 1:10 in Buffer 1.
Detection buffer:	0.1 M Tris, 0.1 M NaCl, pH 9.5

Procedure:

1. After the post-hybridization wash the blots are briefly rinsed (1-5 min.) in the maleate washing buffer with gentle shaking.
2. Then the membranes are incubated for 30 min. in Buffer 2 with gentle shaking.
3. Anti-DIG-AP conjugate (Boehringer Mannheim, Indianapolis, IN, cat. no. 1093274) at 75 mU/ml (1:10,000) in Buffer 2 is used for detection. 75 ml of solution can be used for 3 blots.
4. The membrane is incubated for 30 min. in the antibody solution with gentle shaking.
5. The membrane are washed twice in washing buffer with gentle shaking. About 250 mls is used per wash for 3 blots.
6. The blots are equilibrated for 2-5 min in 60 ml detection buffer.

7. Dilute CSPD (1:200) in detection buffer. (This can be prepared ahead of time and stored in the dark at 4°C).

The following steps must be done individually. Bags (one for detection and one for exposure) are generally cut and ready before doing the following steps.

8. The blot is carefully removed from the detection buffer and excess liquid removed without drying the membrane. The blot is immediately placed in a bag and 1.5 ml of CSPD solution is added. The CSPD solution can be spread over the membrane. Bubbles present at the edge and on the surface of the blot are typically removed by gentle rubbing. The membrane is incubated for 5 min. in CSPD solution.
9. Excess liquid is removed and the membrane is blotted briefly (DNA side up) on Whatman 3MM paper. Do not let the membrane dry completely.
10. Seal the damp membrane in a hybridization bag and incubate for 10 min at 37°C to enhance the luminescent reaction.
11. Expose for 2 hours at room temperature to X-ray film. Multiple exposures can be taken. Luminescence continues for at least 24 hours and signal intensity increases during the first hours.

### EXAMPLE 3: MICROARRAY EXPERIMENTS AND RESULTS

#### **Example 3: MICROARRAY EXPERIMENTS AND RESULTS**

##### **1. Sample Tissue Preparation**

###### **(a) Roots**

Seeds of *Arabidopsis thaliana* (Ws) were sterilized in full strength bleach for less than 5 min., washed more than 3 times in sterile distilled deionized water and plated on MS agar plates. The plates were placed at 4°C for 3 nights and then placed vertically into a growth chamber having 16 hr light/8 hr dark cycles, 23 °C, 70% relative humidity and ~11,000 LUX. After 2 weeks, the roots were cut from the agar, flash frozen in liquid nitrogen and stored at -80°C.(EXPT REP: 108439 and 108434)

**(b) Root Hairless mutants**

Plants mutant at the *rhl* gene locus lack root hairs. This mutation is maintained as a heterozygote.

Seeds of *Arabidopsis thaliana* (Landsberg erecta) mutated at the *rhl* gene locus were sterilized using 30% bleach with 1 ul/ml 20% Triton -X 100 and then vernalized at 4°C for 3 days before being plated onto GM agar plates. Plates were placed in growth chamber with 16 hr light/8 hr. dark, 23°C, 14,500-15,900 LUX, and 70% relative humidity for germination and growth.

After 7 days, seedlings were inspected for root hairs using a dissecting microscope. Mutants were harvested and the cotyledons removed so that only root tissue remained. Tissue was then flash frozen in liquid nitrogen and stored at -80C. (EXPT REP: 108433)

*Arabidopsis thaliana* (Landsberg erecta) seedlings grown and prepared as above were used as controls. (EXPT REP: 108433)

Alternatively, seeds of *Arabidopsis thaliana* (Landsberg erecta), heterozygous for the *rhl1* (root hairless) mutation, were surface-sterilized in 30% bleach containing 0.1% Triton X-100 and further rinsed in sterile water. They were then vernalized at 4°C for 4 days before being plated onto MS agar plates. The plates were maintained in a growth chamber at 24°C with 16 hr light/8 hr dark for germination and growth. After 10 days, seedling roots that expressed the phenotype (i.e. lacking root hairs) were cut below the hypocotyl junction, frozen in liquid nitrogen and stored at -80°C. Those seedlings with the normal root phenotype (heterozygous or wt) were collected as described for the mutant and used as controls.

**(c) Rosette Leaves, Stems, and Siliques**

*Arabidopsis thaliana* (Ws) seed was vernalized at 4°C for 3 days before sowing in Metro-mix soil type 350. Flats were placed in a growth chamber having 16 hr light/8 hr dark, 80% relative humidity, 23°C and 13,000 LUX for germination and growth. After 3 weeks, rosette leaves, stems, and siliques (see EXPT REP: 108436, 108437 and 108438) were harvested, flash frozen in liquid nitrogen and stored at -80°C until use. After 4 weeks, siliques (<5mm, 5-10 mm and >10 mm) were harvested, flash frozen in

liquid nitrogen and stored at -80°C until use. 5 week old whole plants (used as controls) were harvested, flash frozen in liquid nitrogen and kept at -80°C until RNA was isolated.

**(d) Trichomes**

*Arabidopsis thaliana* (Colombia glabrous) inflorescences were used as a control and CS8143 (hairy inflorescence ecotype) inflorescences, having increased trichomes, were used as the experimental sample.

Approximately 10 µl of each type of seed was sown on a flat of 350 soil (containing 0.03% marathon) and vernalized at 4°C for 3 days. Plants were then grown at room temperature under florescent lighting. Young inflorescences were collected at 30 days for the control plants and 37 days for the experimental plants. Each inflorescence was cut into one-half inch (1/2") pieces, flash frozen in liquid nitrogen and stored at -80°C until RNA was isolated.

**(e) Germination**

*Arabidopsis thaliana* seeds (ecotype Ws) were sterilized in bleach and rinsed with sterile water. The seeds were placed in 100mm petri plates containing soaked autoclaved filter paper. Plates were foil-wrapped and left at 4°C for 3 nights to vernalize. After cold treatment, the foil was removed and plates were placed into a growth chamber having 16 hr light/8 hr dark cycles, 23 °C, 70% relative humidity and ~11,000 lux. Seeds were collected 1 d (EXPT REP: 108461), 2 d (EXPT REP: 108462), 3 d (EXPT REP: 108463) and 4 d (EXPT REP: 108464) later, flash frozen in liquid nitrogen and stored at -80°C until RNA was isolated.

**(f) Shoot Apical Meristem**

*Arabidopsis thaliana* (Landsberg erecta) plants mutant at the *stm* gene locus lack shoot meristems, produce aerial rosettes, have a reduced number of flowers per inflorescence, as well as a reduced number of petals, stamens and carpels, and is female sterile. This mutation is maintained as a heterozygote.

Seeds of *Arabidopsis thaliana* (Landsberg erecta) mutated at the *stm* locus were sterilized using 30% bleach with 1 ul/ml 20% Triton -X100. The seeds were vernalized at 4°C for 3 days before being plated onto GM agar plates. Half were then put into a 22°C, 24 hr light growth

chamber and half in a 24°C 16 hr light/8 hr dark growth chamber having 14,500-15,900 LUX, and 70% relative humidity for germination and growth.

After 7 days, seedlings were examined for leaf primordia using a dissecting microscope. Presence of leaf primordia indicated a wild type phenotype. Mutants were selected based on lack of leaf primordia. Mutants were then harvested and hypocotyls removed leaving only tissue in the shoot region. Tissue was then flash frozen in liquid nitrogen and stored at -80°C.

Control tissue was isolated from 5 day old Landsberg erecta seedlings grown in the same manner as above. Tissue from the shoot region was harvested in the same manner as the *stm* tissue, but only contained material from the 24°C, 16 hr light/8 hr dark long day cycle growth chamber. (EXPT REP: 108453)

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 8 days. Seedlings were carefully removed from the sand and the outer layers of leaf sheath removed. About 2 mm sections were cut and flash frozen in liquid nitrogen prior to storage at -80°C. The tissues above the shoot apices (~1 cm long) were cut, treated as above and used as control tissue.

#### **(g) Abscissic acid (ABA)**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at 4°C for 4 days to vernalize. They were then transferred to a growth chamber having grown 16 hr light/8 hr dark, 13,000 LUX, 70% humidity, and 20°C and watered twice a week with 1 L of 1X Hoagland's solution. Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 100 µM ABA in a 0.02% solution of the detergent Silwet L-77. Whole seedlings, including roots, were harvested within a 15 to 20 minute time period at 1 hr and 6 hr after treatment, flash-frozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 100 µM ABA for treatment. Control plants were

treated with water. After 6 hr and 24 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at  $-80^{\circ}\text{C}$ .

**(h) Auxin Responsive**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at  $4^{\circ}\text{C}$  for 4 days to vernalize. They were then transferred to a growth chamber having 16 hr light/8 hr dark, 13,000 LUX, 70% humidity,  $20^{\circ}\text{C}$  and watered twice a week with 1 L of 1X Hoagland's solution (recipe recited in Feldmann et al., (1987) Mol. Gen. Genet. 208: 1-9 and described as complete nutrient solution). Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 100  $\mu\text{M}$  NAA in a 0.02% solution of the detergent Silwet L-77. Aerial tissues (everything above the soil line) were harvested within a 15 to 20 minute time period 1 hr and 6 hrs after treatment, flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light ( $25^{\circ}\text{C}$ )/8 hr dark ( $20^{\circ}\text{C}$ ), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 100  $\mu\text{M}$  NAA for treatment. Control plants were treated with water. After 6 hr and 24 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at  $-80^{\circ}\text{C}$ .

**(i) Cytokinin**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at  $4^{\circ}\text{C}$  for 4 days to vernalize. They were then transferred to a growth chamber having 16 hr light/8 hr dark, 13,000 LUX, 70% humidity,  $20^{\circ}\text{C}$  temperature and watered twice a week with 1 L of 1X Hoagland's solution. Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 100  $\mu\text{M}$  BA in a 0.02% solution of the detergent Silwet L-77. Aerial tissues (everything above the soil line) were harvested within a 15 to 20 minute time period 1 hr and 6 hrs after treatment, flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .



Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 100  $\mu$ M BA for treatment. Control plants were treated with water. After 6 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(j) Brassinosteroid Responsive**

Two separate experiments were performed, one with epi-brassinolide and one with the brassinosteroid biosynthetic inhibitor brassinazole.

In the epi-brassinolide experiments, seeds of wild-type *Arabidopsis thaliana* (ecotype Wassilewskija) and the brassinosteroid biosynthetic mutant *dwf4-1* were sown in trays and left at 4°C for 4 days to vernalize. They were then transferred to a growth chamber having 16 hr light/8 hr dark, 11,000 LUX, 70% humidity and 22°C temperature. Four week old plants were sprayed with a 1  $\mu$ M solution of epi-brassinolide and shoot parts (unopened floral primordia and shoot apical meristems) harvested three hours later. Tissue was flash-frozen in liquid nitrogen and stored at -80°C.(EXPT REP 108480)

In the brassinazole experiments, seeds of wild-type *Arabidopsis thaliana* (ecotype Wassilewskija) were grown as described above. Four week old plants were sprayed with a 1  $\mu$ M solution of brassinazole and shoot parts (unopened floral primordia and shoot apical meristems) harvested three hours later. Tissue was flash-frozen in liquid nitrogen and stored at -80°C.(EXPT REP 108481)

In addition to the spray experiments, tissue was prepared from two different mutants; (1) a *dwf4-1* knock out mutant (EXPT REP: 108478) and (2) a mutant overexpressing the *dwf4-1* gene (EXPT REP: 108479).

Seeds of wild-type *Arabidopsis thaliana* (ecotype Wassilewskija) and of the *dwf4-1* knock out and overexpressor mutants were sown in trays and left at 4°C for 4 days to vernalize. They were then transferred to a growth chamber having 16 hr light/8 hr dark, 11,000 LUX, 70% humidity and 22°C temperature. Tissue from shoot parts (unopened floral primordia and shoot apical meristems) was flash-frozen in liquid nitrogen and stored at -80°C.

Another experiment was completed with seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at 4°C for 4 days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr. dark) conditions, 13,000 LUX light intensity, 70% humidity, 20°C temperature and watered twice a week with 1 L 1X Hoagland's solution(recipe recited in Feldmann et al., (1987) Mol. Gen. Genet. 208: 1-9 and described as complete nutrient solution). Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 0.1 µM Epi-Brassinolite in 0.02% solution of the detergent Silwet L-77. At 1 hr. and 6 hrs. after treatment aerial tissues were harvested within a 15 to 20 minute time period and flash-frozen in liquid nitrogen.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 0.1 µM epi-brassinolide for treatment. Control plants were treated with distilled deionized water. After 24 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

#### **(k) Gibberillic Acid**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at 4°C for 4 days to vernalize. They were then transferred to a growth chamber having 16 hr light/8 hr. dark, 13,000 LUX, 70% humidity, 20°C and watered twice a week with 1 L of 1X Hoagland's solution. Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 100 µM gibberillic acid in a 0.02% solution of the detergent Silwet L-77. At 1 hr. and 6 hrs. after treatment, aerial tissues (everything above the soil line) were harvested within a 15 to 20 minute time period, flash-frozen in liquid nitrogen and stored at -80°C.

Alternatively, seeds of *Arabidopsis thaliana* (ecotype Ws) were sown in Metro-mix soil type 350 and left at 4°C for 3 days to vernalize. They were then transferred to a growth chamber having 16 hr light/8 hr dark, 13,000 LUX, 80% humidity, 20°C temperature and watered every four days with 1.5 L water. 14 days after germination, plants were sprayed with 100 µM

gibberillic acid or with water. Aerial tissues were harvested 1 hr (EXPT REP: 108484), 6 hrs (EXPT REP: 108485), 12 hrs (EXPT REP: 108486), and 24 hrs post-treatment, flash frozen and stored at  $-80^{\circ}\text{C}$ .

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light ( $25^{\circ}\text{C}$ )/8 hr dark ( $20^{\circ}\text{C}$ ), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with  $100\text{ }\mu\text{M}$  gibberillic acid for treatment. Control plants were treated with water. After 1 hr, 6 hr and 12 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at  $-80^{\circ}\text{C}$ .

**(I) Nitrogen: High to Low**

Wild type *Arabidopsis thaliana* seeds (ecotype Ws) were surface sterilized with 30% Clorox, 0.1% Triton X-100 for 5 minutes. Seeds were then rinsed with 4-5 exchanges of sterile double distilled deionized water. Seeds were vernalized at  $4^{\circ}\text{C}$  for 2-4 days in darkness. After cold treatment, seeds were plated on modified 1X MS media (without  $\text{NH}_4\text{NO}_3$  or  $\text{KNO}_3$ ), 0.5% sucrose, 0.5g/L MES pH5.7, 1% phytagar and supplemented with  $\text{KNO}_3$  to a final concentration of 60 mM (high nitrate modified 1X MS media). Plates were then grown for 7 days in a Percival growth chamber at  $22^{\circ}\text{C}$  with 16 hr. light/8 hr dark.

Germinated seedlings were then transferred to a sterile flask containing 50 mL of high nitrate modified 1X MS liquid media. Seedlings were grown with mild shaking for 3 additional days at  $22^{\circ}\text{C}$  in 16 hr. light/8 hr dark (in a Percival growth chamber) on the high nitrate modified 1X MS liquid media.

After three days of growth on high nitrate modified 1X MS liquid media, seedlings were transferred either to a new sterile flask containing 50 mL of high nitrate modified 1X MS liquid media or to low nitrate modified 1X MS liquid media (containing  $20\text{ }\mu\text{M}$   $\text{KNO}_3$ ). Seedlings were grown in these media conditions with mild shaking at  $22^{\circ}\text{C}$  in 16 hr light/ 8 hr dark for the appropriate time points and whole seedlings harvested for total RNA isolation via the Trizol method (LifeTech.). The time points used for the microarray experiments were 10 min. (EXPT REP: 108454) and 1 hour (EXPT REP: 108455) time points for both the high and low nitrate modified 1X MS media.

Alternatively, seeds that were surface sterilized in 30% bleach containing 0.1% Triton X-100 and further rinsed in sterile water, were planted on MS agar, (0.5% sucrose) plates containing 50 mM KNO<sub>3</sub> (potassium nitrate). The seedlings were grown under constant light (3500 LUX) at 22°C. After 12 days, seedlings were transferred to MS agar plates containing either 1mM KNO<sub>3</sub> or 50 mM KNO<sub>3</sub>. Seedlings transferred to agar plates containing 50 mM KNO<sub>3</sub> were treated as controls in the experiment. Seedlings transferred to plates with 1mM KNO<sub>3</sub> were rinsed thoroughly with sterile MS solution containing 1 mM KNO<sub>3</sub>. There were ten plates per transfer. Root tissue was collected and frozen in 15 mL Falcon tubes at various time points which included 1 hour, 2 hours, 3 hours, 4 hours, 6 hours, 9 hours, 12 hours, 16 hours, and 24 hours.

Maize 35A19 Pioneer hybrid seeds were sown on flats containing sand and grown in a Conviron growth chamber at 25°C, 16 hr light/8 hr dark, ~13,000 LUX and 80% relative humidity. Plants were watered every three days with double distilled deionized water. Germinated seedlings are allowed to grow for 10 days and were watered with high nitrate modified 1X MS liquid media (see above). On day 11, young corn seedlings were removed from the sand (with their roots intact) and rinsed briefly in high nitrate modified 1X MS liquid media. The equivalent of half a flat of seedlings were then submerged (up to their roots) in a beaker containing either 500 mL of high or low nitrate modified 1X MS liquid media (see above for details).

At appropriate time points, seedlings were removed from their respective liquid media, the roots separated from the shoots and each tissue type flash frozen in liquid nitrogen and stored at -80°C. This was repeated for each time point. Total RNA was isolated using the Trizol method (see above) with root tissues only.

Corn root tissues isolated at the 4 hr and 16 hr time points were used for the microarray experiments. Both the high and low nitrate modified 1X MS media were used.

**(m) Nitrogen: Low to High**

*Arabidopsis thaliana* ecotype Ws seeds were sown on flats containing 4 L of a 1:2 mixture of Grace Zonolite vermiculite and soil. Flats were watered with 3 L of water and vernalized at 4°C for five days. Flats were placed in a Conviron growth chamber having 16 hr light/8 hr dark at 20°C, 80% humidity and 17,450 LUX. Flats were watered with approximately 1.5 L of water every four days. Mature, bolting plants (24 days after germination) were bottom

treated with 2 L of either a control (100 mM mannitol pH 5.5) or an experimental (50 mM ammonium nitrate, pH 5.5) solution. Roots, leaves and siliques were harvested separately 30, 120 and 240 minutes after treatment, flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

Hybrid maize seed (Pioneer hybrid 35A19) were aerated overnight in deionized water.

Thirty seeds were plated in each flat, which contained 4 liters of Grace zonolite vermiculite. Two liters of water were bottom fed and flats were kept in a Conviron growth chamber with 16 hr light/8 hr dark at  $20^{\circ}\text{C}$  and 80% humidity. Flats were watered with 1 L of tap water every three days. Five day old seedlings were treated as described above with 2 L of either a control (100 mM mannitol pH 6.5) solution or 1 L of an experimental (50 mM ammonium nitrate, pH 6.8) solution. Fifteen shoots per time point per treatment were harvested 10, 90 and 180 minutes after treatment, flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

Alternatively, seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were left at  $4^{\circ}\text{C}$  for 3 days to vernalize. They were then sown on vermiculite in a growth chamber having 16 hours light/8 hours dark, 12,000-14,000 LUX, 70% humidity, and  $20^{\circ}\text{C}$ . They were bottom-watered with tap water, twice weekly. Twenty-four days old plants were sprayed with either water (control) or 0.6% ammonium nitrate at  $4\text{ }\mu\text{L}/\text{cm}^2$  of tray surface. Total shoots and some primary roots were cleaned of vermiculite, flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

#### **(n) Methyl Jasmonate**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at  $4^{\circ}\text{C}$  for 4 days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr. dark, 13,000 LUX, 70% humidity,  $20^{\circ}\text{C}$  temperature and watered twice a week with 1 L of a 1X Hoagland's solution. Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 0.001% methyl jasmonate in a 0.02% solution of the detergent Silwet L-77. At 1 hr and 6 hrs after treatment, whole seedlings, including roots, were harvested within a 15 to 20 minute time period, flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light ( $25^{\circ}\text{C}$ )/8 hr dark ( $20^{\circ}\text{C}$ ), 75% relative humidity and 13,000-14,000 LUX.

Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 0.001% methyl jasmonate for treatment. Control

plants were treated with water. After 24 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at  $-80^{\circ}\text{C}$ .

**(o) Salicylic Acid**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in trays and left at  $4^{\circ}\text{C}$  for 4 days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr dark, 13,000 LUX, 70% humidity,  $20^{\circ}\text{C}$  temperature and watered twice a week with 1 L of a IX Hoagland's solution. Approximately 1,000 14 day old plants were sprayed with 200-250 mls of 5 mM salicylic acid (solubilized in 70% ethanol) in a 0.02% solution of the detergent Silwet L-77. At 1 hr and 6 hrs after treatment, whole seedlings, including roots, were harvested within a 15 to 20 minute time period flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

Alternatively, seeds of wild-type *Arabidopsis thaliana* (ecotype Columbia) and mutant CS3726 were sown in soil type 200 mixed with osmocote fertilizer and Marathon insecticide and left at  $4^{\circ}\text{C}$  for 3 days to vernalize. Flats were incubated at room temperature with continuous light. Sixteen days post germination plants were sprayed with 2 mM SA, 0.02% SilwetL-77 or control solution (0.02% SilwetL-77. Aerial parts or flowers were harvested 1 hr (EXPT REP: 108471 and 108472), 4 hr (EXPT REP: 108469 and 108470), 6 hr (EXPT REP: 108440,) 24 hr (EXPT REP: 108443, 107953 and 107960) and 3 weeks (EXPT REP: 108475, 108476) post-treatment flash frozen and stored at  $-80^{\circ}\text{C}$ .

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light ( $25^{\circ}\text{C}$ )/8 hr dark ( $20^{\circ}\text{C}$ ), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 2 mM SA for treatment. Control plants were treated with water. After 12 hr and 24 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at  $-80^{\circ}\text{C}$ .

**(p) Wounding**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in trays and left at  $4^{\circ}\text{C}$  for three days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr dark, 12,000-14,000 LUX, 70% humidity and  $20^{\circ}\text{C}$ . After 14 days, the leaves were wounded with

forceps. Aerial tissues were harvested 1 hour and 6 hours after wounding. Aerial tissues from unwounded plants served as controls. Tissues were flash-frozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were wounded (one leaf nicked by scissors) and placed in 1-liter beakers of water for treatment. Control plants were treated not wounded. After 1 hr and 6 hr aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(q) Drought stress**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in pots and left at 4°C for three days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr dark, 150,000-160,000 LUX, 20°C and 70% humidity. After 14 days, aerial tissues were cut and left to dry on 3MM Whatman paper in a petri-plate for 1 hour and 6 hours. Aerial tissues exposed for 1 hour and 6 hours to 3 MM Whatman paper wetted with 1X Hoagland's solution served as controls. Tissues were harvested, flash-frozen in liquid nitrogen and stored at -80°C.

Alternatively, *Arabidopsis thaliana* (Ws) seed was vernalized at 4° C for 3 days before sowing in Metromix soil type 350. Flats were placed in a growth chamber with 23°C, 16 hr light/8 hr. dark, 80% relative humidity, ~13,000 LUX for germination and growth. Plants were watered with 1-1.5 L of water every four days. Watering was stopped 16 days after germination for the treated samples, but continued for the control samples. Rosette leaves and stems (EXPT REP 108477, 108482 and 108483), flowers (see EXPT REP: 108473, 108474) and siliques were harvested 2 d, 3 d, 4 d, 5 d, 6 d and 7 d (EXPT REP: 108473) after watering was stopped. Tissue was flash frozen in liquid nitrogen and kept at -80 °C until RNA was isolated. Flowers and siliques were also harvested on day 8 from plants that had undergone a 7 d drought treatment followed by 1 day of watering (EXPT REP: 108474). Control plants (whole plants) were harvested after 5 weeks, flash frozen in liquid nitrogen and stored as above.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber

having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in empty 1-liter beakers at room temperature for treatment. Control plants were placed in water. After 1 hr, 6 hr, 12 hr and 24 hr aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(r) Osmotic stress**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in trays and left at 4°C for three days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr dark, 12,000-14,000 LUX, 20°C, and 70% humidity. After 14 days, the aerial tissues were cut and placed on 3 MM Whatman paper in a petri-plate wetted with 20% PEG (polyethylene glycol-M<sub>r</sub> 8,000) in 1X Hoagland's solution. Aerial tissues on 3 MM Whatman paper containing 1X Hoagland's solution alone served as the control. Aerial tissues were harvested at 1 hour and 6 hours after treatment, flash-frozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 20% PEG (polyethylene glycol-M<sub>r</sub> 8,000) for treatment. Control plants were treated with water. After 1 hr and 6 hr aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 150mM NaCl for treatment. Control plants were treated with water. After 1 hr, 6hr, and 24 hr aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.



**(s) Heat Shock Treatment**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in trays and left at 4°C for three days to vernalize before being transferred to a growth chamber with 16 hr light/8 hr dark, 12,000-14,000 LUX, 70% humidity and 20°C. Fourteen day old plants were transferred to a 42°C growth chamber and aerial tissues were harvested 1 h and 6 h after transfer. Control plants were left at 20°C and aerial tissues were harvested. Tissues were flashfrozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers containing 42°C water for treatment. Control plants were treated with water at 25°C. After 1 hr and 6 hr aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(t) Cold Shock Treatment**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in trays and left at 4°C for three days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr dark, 12,000-14,000 LUX, 20°C and 70% humidity. Fourteen day old plants were transferred to a 4°C dark growth chamber and aerial tissues were harvested 1 hour and 6 hours later. Control plants were maintained at 20°C and covered with foil to avoid exposure to light. Tissues were flash-frozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers containing 4°C water for treatment. Control plants were treated with water at 25°C. After 1 hr and 6 hr aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(u) Oxidative Stress- Hydrogen Peroxide Treatment**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in trays and left at 4°C for three days to vernalize. before being transferred to a growth chamber having 16 hr light/8 hr dark, 12,000-14,000 LUX, 20°C and 70% humidity. Fourteen day old plants were sprayed with 5 mM H<sub>2</sub>O<sub>2</sub> (hydrogen peroxide) in a 0.02% Silwett L-77 solution. Control plants were sprayed with a 0.02% Silwett L-77 solution. Aerial tissues were harvested 1 hour and 6 hours after spraying, flash-frozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 5 mM H<sub>2</sub>O<sub>2</sub> for treatment. Control plants were treated with water. After 1 hr, 6 hr and 24 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(v) Nitric Oxide Treatment**

Seeds of *Arabidopsis thaliana* (Wassilewskija) were sown in trays and left at 4°C for three days to vernalize before being transferred to a growth chamber having 16 hr light/8 hr dark, 12,000-14,000 LUX, 20°C and 70% humidity. Fourteen day old plants were sprayed with 5 mM sodium nitroprusside in a 0.02% Silwett L-77 solution. Control plants were sprayed with a 0.02% Silwett L-77 solution. Aerial tissues were harvested 1 hour and 6 hours after spraying, flash-frozen in liquid nitrogen and stored at -80°C.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 7 days. Seedlings were carefully removed from the sand and placed in 1-liter beakers with 5 mM nitroprusside for treatment. Control plants were treated with water. After 1 hr, 6 hr and 12 hr, aerial and root tissues were separated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(w) S4 Immature Buds, Inflorescence Meristem**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. Inflorescences containing immature floral buds [stages 1-12; Smyth et al., 1990] as well as the inflorescence meristem were harvested and flash frozen in liquid nitrogen.

#### **(x) S5 Flowers (Opened)**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. Mature, unpollinated flowers [stages 12-14; Smyth et al. 1990] were harvested and flash frozen in liquid nitrogen.

#### **(y) S6 Siliques (All Stages)**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. Siliques bearing developing seeds containing post fertilization through pre-heart stage [0-72 hours after fertilization (HAF)], heart- through early curled cotyledon stage [72-120 HAF] and late-curved cotyledon stage [≥120 HAF] embryos were harvested separately and pooled prior to RNA isolation in a mass ratio of 1:1:1. The tissues were then flash frozen in liquid nitrogen. Description of the stages of *Arabidopsis* embryogenesis used were reviewed by Bowman (1994).

#### **(z) ARABIDOPSIS ENDOSPERM**

##### **mea/mea Fruits 0-10 mm**

Seeds of *Arabidopsis thaliana* heterozygous for the *fertilization- independent endosperm1 (fie1)* [Ohad et al., 1996; ecotype Landsberg *erecta* (Ler)] were sown in pots and left at 4°C for two to three days to vernalize. Kiyosue et al. (1999) subsequently determined that

*fiel* was allelic to the gametophytic maternal effect mutant *medea* (Grossniklaus et al., 1998). Imbibed seeds were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 1-2 siliques (fruits) bearing developing seeds just prior to dessication [9 days after flowering (DAF)] were selected from each plant and were hand-dissected to identify wild-type, *mea/+* heterozygotes, and *mea/mea* homozygous mutant plants. At this stage, homozygous *mea/mea* plants produce short siliques that contain >70% aborted seed and can be distinguished from those produced by wild-type (100% viable seed) and *mea/+* heterozygous (50% viable seed) plants (Ohad et al., 1996; Grossniklaus et al., 1998; Kiyosue et al., 1999). Siliques 0-10 mm in length containing developing seeds 0-9 DAF produced by homozygous *mea/mea* plants were harvested and flash frozen in liquid nitrogen.

#### **Pods 0-10 mm (Control Tissue for Sample 70)**

Seeds of *Arabidopsis thaliana* heterozygous for the fertilization- independent *endosperm1 (fiel)* [Ohad et al., 1996; ecotype Landsberg *erecta* (Ler)] were sown in pots and left at 4°C for two to three days to vernalize. Kiyosue et al. (1999) subsequently determined that *fiel* was allelic to the gametophytic maternal effect mutant *medea* (Grossniklaus et al., 1998). Imbibed seeds were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 1-2 siliques (fruits) bearing developing seeds just prior to dessication [9 days after flowering (DAF)] were selected from each plant and were hand-dissected to identify wild-type, *mea/+* heterozygotes, and *mea/mea* homozygous mutant plants. At this stage, homozygous *mea/mea* plants produce short siliques that contain >70% aborted seed and can be distinguished from those produced by wild-type (100% viable seed) and *mea/+* heterozygous (50% viable seed) plants (Ohad et al., 1996; Grossniklaus et al., 1998; Kiyosue et al., 1999). Siliques 0-10 mm in length containing developing seeds 0-9 DAF produced by segregating wild-type plants were opened and the seeds removed. The remaining tissues (pods minus seed) were harvested and flash frozen in liquid nitrogen.

#### **(aa) ARABIDOPSIS SEEDS**

**Fruits (pod + seed) 0-5 mm**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were represented by the enclosed embryos. Description of the stages of *Arabidopsis* embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Siliques 0-5 mm in length containing post fertilization through pre-heart stage [0-72 hours after fertilization (HAF)] embryos were harvested and flash frozen in liquid nitrogen.

**Fruits(pod + seed) 5-10 mm**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were represented by the enclosed embryos. Description of the stages of *Arabidopsis* embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Siliques 5-10 mm in length containing heart- through early upturned-U- stage [72-120 hours after fertilization (HAF)] embryos were harvested and flash frozen in liquid nitrogen.

**Fruits(pod + seed) >10 mm**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were

represented by the enclosed embryos. Description of the stages of Arabidopsis embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Siliques >10 mm in length containing green, late upturned-U- stage [>120 hours after fertilization (HAF)-9 days after flowering (DAF)] embryos were harvested and flash frozen in liquid nitrogen.

#### **Green Pods 5-10 mm (Control Tissue for Samples 72-74)**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were represented by the enclosed embryos. Description of the stages of Arabidopsis embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Green siliques 5-10 mm in length containing developing seeds 72-120 hours after fertilization (HAF)] were opened and the seeds removed. The remaining tissues (green pods minus seed) were harvested and flash frozen in liquid nitrogen.

#### **Green Seeds from Fruits >10 mm**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were represented by the enclosed embryos. Description of the stages of Arabidopsis embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Green siliques >10

mm in length containing developing seeds up to 9 days after flowering (DAF)] were opened and the seeds removed and harvested and flash frozen in liquid nitrogen.

**Brown Seeds from Fruits >10 mm**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were represented by the enclosed embryos. Description of the stages of *Arabidopsis* embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Yellowing siliques >10 mm in length containing brown, desiccating seeds >11 days after flowering (DAF)] were opened and the seeds removed and harvested and flash frozen in liquid nitrogen.

**Green/Brown Seeds from Fruits >10 mm**

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature. 3-4 siliques (fruits) bearing developing seeds were selected from at least 3 plants and were hand-dissected to determine what developmental stage(s) were represented by the enclosed embryos. Description of the stages of *Arabidopsis* embryogenesis used in this determination were summarized by Bowman (1994). Silique lengths were then determined and used as an approximate determinant for embryonic stage. Green siliques >10 mm in length containing both green and brown seeds >9 days after flowering (DAF)] were opened and the seeds removed and harvested and flash frozen in liquid nitrogen.

### Mature Seeds (24 hours after imbibition)

Mature dry seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown onto moistened filter paper and left at 4°C for two to three days to vernalize. Imbibed seeds were then transferred to a growth chamber [16 hr light: 8 hr dark conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature], the emerging seedlings harvested after 48 hours and flash frozen in liquid nitrogen.

### Mature Seeds (Dry)

Seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 70% humidity, and 22°C temperature and taken to maturity. Mature dry seeds are collected, dried for one week at 28°C, and vernalized for one week at 4°C before used as a source of RNA.

### Ovules

Seeds of *Arabidopsis thaliana* heterozygous for *pistillata* (*pi*) [ecotype Landsberg *erecta* (*Ler*)] were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light: 8 hr dark) conditions, 7000-8000 LUX light intensity, 76% humidity, and 24°C temperature. Inflorescences were harvested from seedlings about 40 days old. The inflorescences were cut into small pieces and incubated in the following enzyme solution (pH 5) at room temperature for 0.5-1 hr.: 0.2% pectolyase Y-23, 0.04% pectinase, 5 mM MES, 3% Sucrose and MS salts (1900 mg/l KNO<sub>3</sub>, 1650 mg/l NH<sub>4</sub>NO<sub>3</sub>, 370 mg/l MgSO<sub>4</sub> • 7 H<sub>2</sub>O, 170 mg/l KH<sub>2</sub>PO<sub>4</sub>, 440 mg/l CaCl<sub>2</sub> • 2 H<sub>2</sub>O, 6.2 mg/l H<sub>2</sub>BO<sub>3</sub>, 15.6 mg/l MnSO<sub>4</sub> • 4 H<sub>2</sub>O, 8.6 mg/l ZnSO<sub>4</sub> • 7 H<sub>2</sub>O, 0.25 mg/l NaMoO<sub>4</sub> • 2 H<sub>2</sub>O, 0.025 mg/l CuCO<sub>3</sub> • 5 H<sub>2</sub>O, 0.025 mg/l CoCl<sub>2</sub> • 6 H<sub>2</sub>O, 0.83 mg/l KI, 27.8 mg/l FeSO<sub>4</sub> • 7 H<sub>2</sub>O, 37.3 mg/l Disodium EDTA, pH 5.8). At the end of the incubation the mixture of inflorescence material and enzyme solution was passed through a size 60 sieve and then through a sieve with a pore size of 125 µm. Ovules greater than 125 µm in diameter were collected, rinsed twice in B5 liquid medium (2500 mg/l KNO<sub>3</sub>, 250 mg/l MgSO<sub>4</sub> • 7 H<sub>2</sub>O, 150 mg/l NaH<sub>2</sub>PO<sub>4</sub> • H<sub>2</sub>O, 150 mg/l CaCl<sub>2</sub> • 2 H<sub>2</sub>O, 134 mg/l (NH<sub>4</sub>)<sub>2</sub> CaCl<sub>2</sub> • SO<sub>4</sub>, 3 mg/l H<sub>2</sub>BO<sub>3</sub>, 10



mg/l  $\text{MnSO}_4 \bullet 4 \text{H}_2\text{O}$ , 2  $\text{ZnSO}_4 \bullet 7 \text{H}_2\text{O}$ , 0.25 mg/l  $\text{NaMoO}_4 \bullet 2 \text{H}_2\text{O}$ , 0.025 mg/l  $\text{CuCO}_3 \bullet 5 \text{H}_2\text{O}$ , 0.025 mg/l  $\text{CoCl}_2 \bullet 6 \text{H}_2\text{O}$ , 0.75 mg/l KI, 40 mg/l EDTA sodium ferric salt, 20 g/l sucrose, 10 mg/l Thiamine hydrochloride, 1 mg/l Pyridoxine hydrochloride, 1 mg/l Nicotinic acid, 100 mg/l myo-inositol, pH 5.5)), rinsed once in deionized water and flash frozen in liquid nitrogen.

5 The supernatant from the 125  $\mu\text{m}$  sieving was passed through subsequent sieves of 50  $\mu\text{m}$  and 32  $\mu\text{m}$ . The tissue retained in the 32  $\mu\text{m}$  sieve was collected and mRNA prepared for use as a control.

10 **(bb) Herbicide treatment**

*Arabidopsis thaliana* (Ws) seeds were sterilized for 5 min. with 30% bleach, 50  $\mu\text{l}$  Triton in a total volume of 50 ml. Seeds were vernalized at 4°C for 3 days before being plated onto GM agar plates at a density of about 144 seeds per plate. Plates were incubated in a Percival growth chamber having 16 hr light/8 hr dark, 80% relative humidity, 22 °C and 11,000 LUX for 14 days.

Plates were sprayed (~0.5 mls/plate) with water, Finale (1.128 g/L), Glean (1.88 g/L), RoundUp (0.01 g/L) or Trimec (0.08 g/L). Tissue was collected and flash frozen in liquid nitrogen at the following time points: 0, 1, 2, 4 (EXPT REP: 107871 (Finale), 107881 (Glean), 107896 (Round-up) and 107886 (Trimec)), 8, 12 (EXPT REP: 108467 (Finale), 108468 (Glean), 108465 (Round-up) and 108466, 107891 (Trimec)), and 24 hours. Frozen tissue was stored at -80°C prior to RNA isolation.

20 **(cc) Ap2**

Seeds of *Arabidopsis thaliana* (ecotype Landsberg erecta) and floral mutant *apetala2* (Jofuku et al., 1994, Plant Cell 6:1211-1225) were sown in pots and left at 4°C for two to three days to vernalize. They were then transferred to a growth chamber. Plants were grown under long-day (16 hr light, 8 hr dark) conditions 7000-8000 LUX light intensity, 70% humidity and 22 °C temperature. Inflorescences containing immature floral buds (stages 1-7; Bowman, 1994) as well as the inflorescence meristem were harvested and flashfrozen. Polysomal polyA+ RNA was isolated from tissue according to Cox and Goldberg, 1988).

30 **(dd) Protein Degradation**

*Arabidopsis thaliana* (ecotype Ws) wild-type and 13B12-1 (homozygous) mutant seed were sown in pots containing Metro-mix 350 soil and incubated at 4°C for four days. Vernalized seeds were germinated in the greenhouse (16 hr light/8 hr dark) over a 7 day period. Mutant seedlings were sprayed with 0.02% (active ingredient) Finale to confirm their transgenic standing. Plants were grown until the mutant phenotype (either multiple pistils in a single flower and/or multiple branching per node) was apparent. Young inflorescences immediately forming from the multiple-branched stems were cut and flash frozen in liquid nitrogen. Young inflorescences from wild-type plants grown in parallel and under identical conditions were collected as controls. All collected tissue was stored at -80°C until RNA isolation. (EXPT REP 108451)

**(ee) Root tips**

Seeds of *Arabidopsis thaliana* (ecotype Ws) were placed on MS plates and vernalized at 4°C for 3 days before being placed in a 25°C growth chamber having 16 hr light/8 hr dark, 70% relative humidity and about 3 W/m<sup>2</sup>. After 6 days, young seedlings were transferred to flasks containing B5 liquid medium, 1% sucrose and 0.05 mg/l indole-3-butyric acid. Flasks were incubated at room temperature with 100 rpm agitation. Media was replaced weekly. After three weeks, roots were harvested and incubated for 1 hr with 2% pectinase, 0.2% cellulase, pH 7 before straining through a #80 (Sigma) sieve. The root body material remaining on the sieve (used as the control) was flash frozen and stored at -80°C until use. The material that passed through the #80 sieve was strained through a #200 (Sigma) sieve and the material remaining on the sieve (root tips) was flash frozen and stored at -80°C until use. Approximately 10 mg of root tips were collected from one flask of root culture.

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. Covered flats were watered every three days for 8 days. Seedlings were carefully removed from the sand and the root tips (~2 mm long) were removed and flash frozen in liquid nitrogen prior to storage at -80°C. The tissues above the root tips (~1 cm long) were cut, treated as above and used as control tissue.

**(ff) rt1**

The *rt1* allele is a variation of *rt1 rootless1* and is recessive. Plants displaying the *rt1* phenotype have few or no secondary roots.

Seed from plants segregating for *rt1* were sown on sand and placed in a growth chamber having 16 hr light/8 hr dark, 13,000 LUX, 70% humidity and 20°C temperature. Plants were watered every three days with tap water. Eleven (11) day old seedlings were carefully removed from the sand, keeping the roots intact. *rt1*-type seedlings were separated from their wild-type counterparts and the root tissue isolated. Root tissue from normal seedlings (control) and *rt1* mutants were flash frozen in liquid nitrogen and stored at -80°C until use.

**(gg) Imbibed seed**

Seeds of maize hybrid 35A (Pioneer) were sown in water-moistened sand in covered flats (10 rows, 5-6 seed/row) and covered with clear, plastic lids before being placed in a growth chamber having 16 hr light (25°C)/8 hr dark (20°C), 75% relative humidity and 13,000-14,000 LUX. One day after sowing, whole seeds were flash frozen in liquid nitrogen prior to storage at -80°C. Two days after sowing, embryos and endosperm were isolated and flash frozen in liquid nitrogen prior to storage at -80°C. On days 3-6, aerial tissues, roots and endosperm were isolated and flash frozen in liquid nitrogen prior to storage at -80°C.

**(hh) Rough Sheath2-R (rs2-R) Mutants (1400-6/S-17)**

This experiment was conducted to identify abnormally expressed genes in the shoot apex of *rough sheath2-R (rs2-R)* mutant plants. *rs2* encodes a myb domain DNA binding protein that functions in repression of several shoot apical meristem expressed homeobox genes. Two homeobox gene targets are known for *rs2* repression, *rough sheath1*, *liguleless 3*. The recessive loss of function phenotype of *rs2-R* homozygous plants is described in Schneeberger et al. 1998 Development 125: 2857-2865.

The seed stock genetically segregates 1:1 for *rs2-R/rs2-R* : *rs2-R/+*

Preparation of tissue samples: 160 seedlings pooled from 2 and 3 week old plants grown in sand. Growth conditions; Conviron #107 @ 12 hr days/12hr night, 25°C, 75% humidity. Shoot apex was dissected to include leaf three and older. (Pictures available upon request).

- 1) *rough sheath2-R* homozygous (mutant) shoot apex
- 2) *rough sheath2-R* heterozygous (wt, control) shoot apex

**(ii) Leaf Mutant 3642:**

Mutant 3642 is a recessive mutation that causes abnormal leaf development. The leaves of mutant 3642 plants are characterized by leaf twisting and irregular leaf shape. Mutant 3642 plants also exhibit abnormally shaped floral organs which results in reduced fertility.

Seed segregating for the mutant phenotype was sown in Metro-mix 350 soil and grown in a Conviron growth chamber with watering by sub-irrigation twice a week.

Environmental conditions were set at 20 degrees Celsius, 70% humidity with an 8 hour day, 16 hour night light regime. Plants were harvested after 4 weeks of growth and the entire aerial portion of the plant was harvested and immediately frozen in liquid nitrogen and stored at -80C. Mutant phenotype plants were harvested separately from normal phenotype plants, which serve as the control tissue.

**(jj) Flowers (green, white or buds)**

Approximately 10 µl of *Arabidopsis thaliana* seeds (ecotype Ws) were sown on 350 soil (containing 0.03% marathon) and vernalized at 4C for 3 days. Plants were then grown at room temperature under fluorescent lighting until flowering. Flowers were harvested after 28 days in three different categories. Buds that had not opened at all and were completely green were categorized as "flower buds" (also referred to as green buds by the investigator). Buds that had started to open, with white petals emerging slightly were categorized as "green flowers" (also referred to as white buds by the investigator). Flowers that had opened mostly (with no silique elongation) with white petals completely visible were categorized as "white flowers" (also referred to as open flowers by the investigator). Buds and flowers were harvested with forceps, flash frozen in liquid nitrogen and stored at -80C until RNA was isolated.

**2. Microarray Hybridization Procedures**

Microarray technology provides the ability to monitor mRNA transcript levels of thousands of genes in a single experiment. These experiments simultaneously hybridize two

differentially labeled fluorescent cDNA pools to glass slides that have been previously spotted with cDNA clones of the same species. Each arrayed cDNA spot will have a corresponding ratio of fluorescence that represents the level of disparity between the respective mRNA species in the two sample pools. Thousands of polynucleotides can be spotted on one slide, and each experiment generates a global expression pattern.

#### COATING SLIDES

The microarray consists of a chemically coated microscope slide, referred herein as a "chip" with numerous polynucleotide samples arrayed at a high density. The poly-L-lysine coating allows for this spotting at high density by providing a hydrophobic surface, reducing the spreading of spots of DNA solution arrayed on the slides. Glass microscope slides (Gold Seal #3010 manufactured by Gold Seal Products, Portsmouth, New Hampshire, USA) were coated with a 0.1%W/V solution of Poly-L-lysine (Sigma, St. Louis, Missouri) using the following protocol:

Slides were placed in slide racks (Shandon Lipshaw #121). The racks were then put in chambers (Shandon Lipshaw #121).

Cleaning solution was prepared:

70 g NaOH was dissolved in 280 mL ddH<sub>2</sub>O.

420 mL 95% ethanol was added. The total volume was 700 mL (= 2 X 350 mL); it was stirred until completely mixed.

If the solution remained cloudy, ddH<sub>2</sub>O was added until clear.

3. The solution was poured into chambers with slides; the chambers were covered with glass lids. The solution was mixed on an orbital shaker for 2 hr.

4. The racks were quickly transferred to fresh chambers filled with ddH<sub>2</sub>O. They were rinsed vigorously by plunging racks up and down.

Rinses were repeated 4X with fresh ddH<sub>2</sub>O each time, to remove all traces of NaOH-ethanol.

5. Polylysine solution was prepared:

70 mL poly-L-lysine + 70 mL tissue culture PBS in 560 mL water, using plastic graduated cylinder and beaker.

6. Slides were transferred to polylysine solution and shaken for 1 hr.

7. The rack was transferred to a fresh chambers filled with ddH<sub>2</sub>O. It was plunged up and down 5X to rinse.
8. The slides were centrifuged on microtiter plate carriers (paper towels were placed below the rack to absorb liquid) for 5 min. @ 500 rpm. The slide racks were transferred to empty chambers with covers.
9. Slide racks were dried in a 45C oven for 10 min.
10. The slides were stored in a closed plastic slide box.
11. Normally, the surface of lysine coated slides was not very hydrophobic immediately after this process, but became increasingly hydrophobic with storage. A hydrophobic surface helped ensure that spots didn't run together while printing at high densities. After they aged for 10 days to a month the slides were ready to use. However, coated slides that have been sitting around for long periods of time were usually too old to be used. This was because they developed opaque patches, visible when held to the light, and these resulted in high background hybridization from the fluorescent probe.

Alternatively, precoated glass slides were purchased from TeleChem International, Inc. (Sunnyvale, CA, 94089; catalog number SMM-25, Superamine substrates).

#### PCR AMPLIFICATION OF CDNA CLONE INSERTS

Polynucleotides were amplified from Arabidopsis cDNA clones using insert specific probes. The resulting 100uL PCR reactions were purified with Qiaquick 96 PCR purification columns (Qiagen, Valencia, California, USA) and eluted in 30 uL of 5mM Tris. 8.5uL of the elution were mixed with 1.5uL of 20X SSC to give a final spotting solution of DNA in 3X SSC. The concentrations of DNA generated from each clone varied between 10-100 ng/ul, but were usually about 50 ng/ul.

#### ARRAYING OF PCR PRODUCTS ON GLASS SLIDES

PCR products from cDNA clones were spotted onto the poly-L-Lysine coated glass slides using an arrangement of quill-tip pins (ChipMaker 3 spotting pins; Telechem, International, Inc., Sunnyvale, California, USA) and a robotic arrayer (PixSys 3500, Cartesian Technologies, Irvine, California, USA). Around 0.5 nl of a prepared PCR product was spotted at each location to produce spots with approximately 100um diameters. Spot center-to-center spacing was from 180

um to 210um depending on the array. Printing was conducted in a chamber with relative humidity set at 50%.

Slides containing maize sequences were purchased from Agilent Technology (Palo Alto, CA 94304).

#### POST-PROCESSING OF SLIDES

After arraying, slides were processed through a series of steps – rehydration, UV cross-linking, blocking and denaturation - required prior to hybridization. Slides were rehydrated by placing them over a beaker of warm water (DNA face down), for 2-3 sec, to distribute the DNA more evenly within the spots, and then snap dried on a hot plate (DNA side, face up). The DNA was then cross-linked to the slides by UV irradiation (60-65mJ; 2400 Stratalinker, Stratagene, La Jolla, California, USA).

Following this a blocking step was performed to modify remaining free lysine groups, and hence minimize their ability to bind labeled probe DNA. To achieve this the arrays were placed in a slide rack. An empty slide chamber was left ready on an orbital shaker. The rack was bent slightly inwards in the middle, to ensure the slides would not run into each other while shaking. The blocking solution was prepared as follows:

3x 350-ml glass chambers (with metal tops) were set to one side, and a large round Pyrex dish with dH<sub>2</sub>O was placed ready in the microwave. At this time, 15ml sodium borate was prepared in a 50 ml conical tube.

6-g succinic anhydride was dissolved in approx. 325-350 mL 1-methyl-2-pyrrolidinone. Rapid addition of reagent was crucial.

- a. Immediately after the last flake of the succinic anhydride dissolved, the 15-mL sodium borate was added.
- b. Immediately after the sodium borate solution mixed in, the solution was poured into an empty slide chamber.
- c. The slide rack was plunged rapidly and evenly in the solution. It was vigorously shaken up and down for a few seconds, making sure slides never left the solution.
- d. It was mixed on an orbital shaker for 15-20 min. Meanwhile, the water in the Pyrex dish (enough to cover slide rack) was heated to boiling.

Following this, the slide rack was gently plunge in the 95C water (just stopped boiling) for 2 min. Then the slide rack was plunged 5X in 95% ethanol. The slides and rack were centrifuged for 5 min. @ 500 rpm. The slides were loaded quickly and evenly onto the carriers to avoid streaking. The arrays were used immediately or store in slide box.

The Hybridization process began with the isolation of mRNA from the two tissues (see "*Isolation of total RNA*" and "*Isolation of mRNA*", below) in question followed by their conversion to single stranded cDNA (see "*Generation of probes for hybridization*", below). The cDNA from each tissue was independently labeled with a different fluorescent dye and then both samples were pooled together. This final differentially labeled cDNA pool was then placed on a processed microarray and allowed to hybridize (see "*Hybridization and wash conditions*", below).

#### ISOLATION OF TOTAL RNA

Approximately 1 g of plant tissue was ground in liquid nitrogen to a fine powder and transferred into a 50-ml centrifuge tube containing 10 ml of Trizol reagent. The tube was vigorously vortexed for 1 min and then incubated at room temperature for 10-20 min. on an orbital shaker at 220 rpm. Two ml of chloroform was added to the tube and the solution vortexed vigorously for at least 30-sec before again incubating at room temperature with shaking. The sample was then centrifuged at 12,000 X g (10,000 rpm) for 15-20 min at 4°C. The aqueous layer was removed and mixed by inversion with 2.5 ml of 1.2 M NaCl/0.8 M Sodium Citrate and 2.5 ml of isopropyl alcohol added. After a 10 min. incubation at room temperature, the sample was centrifuged at 12,000 X g (10,000 rpm) for 15 min at 4°C. The pellet was washed with 70% ethanol, re-centrifuged at 8,000 rpm for 5 min and then air dried at room temperature for 10 min. The resulting total RNA was dissolved in either TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) or DEPC (diethylpyrocarbonate) treated deionized water (RNase-free water). For subsequent isolation of mRNA using the Qiagen kit, the total RNA pellet was dissolved in RNase-free water.

#### ISOLATION OF mRNA



mRNA was isolated using the Qiagen Oligotex mRNA Spin-Column protocol (Qiagen, Valencia, California). Briefly, 500  $\mu$ l OBB buffer (20 mM Tris-Cl, pH 7.5, 1 M NaCl, 2 mM EDTA, 0.2% SDS) was added to 500  $\mu$ l of total RNA (0.5 – 0.75 mg) and mixed thoroughly. The sample was first incubated at 70°C for 3 min, then at room temperature for 10 minutes and finally centrifuged for 2 min at 14,000 – 18,000 X g. The pellet was resuspended in 400  $\mu$ l OW2 buffer (10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 1 mM EDTA) by vortexing, the resulting solution placed on a small spin column in a 1.5 ml RNase-free microcentrifuge tube and centrifuged for 1 min at 14,000 – 18,000 X g. The spin column was transferred to a new 1.5 ml RNase-free microcentrifuge tube and washed with 400  $\mu$ l of OW2 buffer. To release the isolated mRNA from the resin, the spin column was again transferred to a new RNase-free 1.5 ml microcentrifuge tube, 20-100  $\mu$ l 70°C OEB buffer (5 mM Tris-Cl, pH 7.5) added and the resin resuspended in the resulting solution via pipeting. The mRNA solution was collected after centrifuging for 1 min at 14,000 – 18,000 X g.

Alternatively, mRNA was isolated using the Stratagene Poly(A) Quik mRNA Isolation Kit (Stratagene, La Jolla, California). Here, up to 0.5 mg of total RNA (maximum volume of 1 ml) was incubated at 65°C for 5 minutes, snap cooled on ice and 0.1X volumes of 10X sample buffer (10mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0) 5 M NaCl) added. The RNA sample was applied to a prepared push column and passed through the column at a rate of ~1 drop every 2 sec. The solution collected was reapplied to the column and collected as above. 200  $\mu$ l of high salt buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.5 NaCl) was applied to the column and passed through the column at a rate of ~1 drop every 2 sec. This step was repeated and followed by three low salt buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.1 M NaCl) washes performed in a similar manner. mRNA was eluted by applying to the column four separate 200  $\mu$ l aliquots of elution buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA) preheated to 65°C. Here, the elution buffer was passed through the column at a rate of 1 drop/sec. The resulting mRNA solution was precipitated by adding 0.1X volumes of 10X sample buffer, 2.5 volumes of ice-cold 100% ethanol, incubating overnight at -20°C and centrifuging at 14,000-18,000 X g for 20-30 min at 4°C. The pellet was washed with 70% ethanol and air dried for 10 min. at room temperature before resuspension in RNase-free deionized water.

#### PREPARATION OF YEAST CONTROLS

Plasmid DNA was isolated from the following yeast clones using Qiagen filtered maxiprep kits (Qiagen, Valencia, California): YAL022c(Fun26), YAL031c(Fun21), YBR032w, YDL131w, YDL182w, YDL194w, YDL196w, YDR050c and YDR116c. Plasmid DNA was linearized with either *Bsr*BI (YAL022c(Fun26), YAL031c(Fun21), YDL131w, YDL182w, YDL194w, YDL196w, YDR050c) or *Afl*III (YBR032w, YDR116c) and isolated.

#### *In Vitro Transcription of Yeast Clones*

The following solution was incubated at 37°C for 2 hours: 17 µl of isolated yeast insert DNA (1 µg), 20 µl 5X buffer, 10 µl 100 mM DTT, 2.5 µl (100 U) RNasin, 20 µl 2.5 mM (ea.) rNTPs, 2.7 µl (40U) SP6 polymerase and 27.8 µl RNase-free deionized water. 2 µl (2 U) Ampli DNase I was added and the incubation continued for another 15 min. 10 µl 5M NH<sub>4</sub>OAC and 100 µl phenol:chloroform:isoamyl alcohol (25:24:1) were added, the solution vortexed and then centrifuged to separate the phases. To precipitate the RNA, 250 µl ethanol was added and the solution incubated at -20°C for at least one hour. The sample was then centrifuged for 20 min at 4°C at 14,000-18,000 X g, the pellet washed with 500 µl of 70% ethanol, air dried at room temperature for 10 min and resuspended in 100 µl of RNase-free deionized water. The precipitation procedure was then repeated.

Alternatively, after the two-hour incubation, the solution was extracted with phenol:chloroform once before adding 0.1 volume 3M sodium acetate and 2.5 volumes of 100% ethanol. The solution was centrifuged at 15,000rpm, 4°C for 20 minutes and the pellet resuspended in RNase-free deionized water. The DNase I treatment was carried out at 37°C for 30 minutes using 2 U of Ampli DNase I in the following reaction condition: 50 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>. The DNase I reaction was then stopped with the addition of NH<sub>4</sub>OAC and phenol:chloroform:isoamyl alcohol (25:24:1), and RNA isolated as described above.

0.15-2.5 ng of the *in vitro* transcript RNA from each yeast clone were added to each plant mRNA sample prior to labeling to serve as positive (internal) probe controls.

#### *GENERATION OF PROBES FOR HYBRIDIZATION*

##### *Generation of labeled probes for hybridization from first-strand cDNA*

Hybridization probes were generated from isolated mRNA using an Atlas<sup>TM</sup> Glass Fluorescent Labeling Kit (Clontech Laboratories, Inc., Palo Alto, California, USA). This entails

a two step labeling procedure that first incorporates primary aliphatic amino groups during cDNA synthesis and then couples fluorescent dye to the cDNA by reaction with the amino functional groups. Briefly, 5 µg of oligo(dT)<sub>18</sub> primer d(TTTTTTTTTTTTTTTTTTV) was mixed with Poly A+ mRNA (1.5 - 2 µg mRNA isolated using the Qiagen Oligotex mRNA Spin-  
5 Column protocol or the Stratagene Poly(A) Quik mRNA Isolation protocol (Stratagene, La Jolla, California, USA)) in a total volume of 25 µl. The sample was incubated in a thermocycler at 70°C for 5 min, cooled to 48°C and 10 µl of 5X cDNA Synthesis Buffer (kit supplied), 5 µl 10X dNTP mix (dATP, dCTP, dGTP, dTTP and aminoallyl-dUTP; kit supplied), 7.5 µl deionized water and 2.5 µl MMLV Reverse Transcriptase (500U) added. The reaction was then incubated at 48°C for 30 minutes, followed by 1hr incubation at 42°C. At the end of the incubation the reaction was heated to 70°C for 10 min, cooled to 37°C and 0.5 µl (5 U) RNase H added, before incubating for 15 min at 37°C. The solution was vortexed for 1 min after the addition of 0.5 µl 0.5 M EDTA and 5 µl of QuickClean Resin (kit supplied) then centrifuged at 14,000-18,000 X g for 1 min. After removing the supernatant to a 0.45 µm spin filter (kit supplied), the sample was again centrifuged at 14,000-18,000 X g for 1 min, and 5.5 µl 3 M sodium acetate and 137.5 µl of 100% ethanol added to the sample before incubating at -20°C for at least 1 hr. The sample was then centrifuged at 14,000-18,000 X g at 4°C for 20 min, the resulting pellet washed with 500 µl 70% ethanol, air-dried at room temperature for 10 min and resuspended in 10 µl of 2X fluorescent labeling buffer (kit provided). 10 µl each of the fluorescent dyes Cy3 and Cy5  
15 (Amersham Pharmacia (Piscataway, New Jersey, USA); prepared according to Atlas™ kit directions of Clontech) were added and the sample incubated in the dark at room temperature for 30 min.

The fluorescently labeled first strand cDNA was precipitated by adding 2 µl 3M sodium acetate and 50 µl 100% ethanol, incubated at -20°C for at least 2 hrs, centrifuged at 14,000-  
25 18,000 X g for 20 min, washed with 70% ethanol, air-dried for 10 min and dissolved in 100 µl of water.

Alternatively, 3-4 µg mRNA, 2.5 (~8.9 ng of in vitro translated mRNA) µl yeast control and 3 µg oligo dTV (TTTTTTTTTTTTTTTTT(A/C/G); Sequence ID No.: X) were mixed in a total volume of 24.7 µl. The sample was incubated in a thermocycler at 70°C for 10 min. before  
30 chilling on ice. To this, 8 µl of 5X first strand buffer (SuperScript II RNase H- Reverse

Transcriptase kit from Invitrogen (Carlsbad, California 92008); cat no. 18064022), 0.8 °C of aa-dUTP/dNTP mix (50X; 25mM dATP, 25mM dGTP, 25mM dCTP, 15mM dTTP, 10mM aminoallyl-dUTP), 4 µl of 0.1 M DTT and 2.5 µl (500 units) of Superscript R.T.II enzyme (Stratagene) were added. The sample was incubated at 42°C for 2 hours before a mixture of 10 °C of 1M NaOH and 10°C of 0.5 M EDTA were added. After a 15 minute incubation at 65°C, 25 µl of 1 M Tris pH 7.4 was added. This was mixed with 450 µl of water in a Microcon 30 column before centrifugation at 11,000 X g for 12 min. The column was washed twice with 450 µl (centrifugation at 11,000 g, 12 min.) before eluting the sample by inverting the Microcon column and centrifuging at 11,000 X g for 20 seconds. Sample was dehydrated by centrifugation under vacuum and stored at -20°C.

Each reaction pellet was dissolved in 9 µl of 0.1 M carbonate buffer (0.1M sodium carbonate and sodium bicarbonate, pH=8.5-9) and 4.5 µl of this placed in two microfuge tubes. 4.5 µl of each dye (in DMSO) were added and the mixture incubated in the dark for 1 hour. 4.5 µl of 4 M hydroxylamine was added and again incubated in the dark for 15 minutes.

Regardless of the method used for probe generation, the probe was purified using a Qiagen PCR cleanup kit (Qiagen, Valencia, California, USA), and eluted with 100 µl EB (kit provided). The sample was loaded on a Microcon YM-30 (Millipore, Bedford, Massachusetts, USA) spin column and concentrated to 4-5 µl in volume. Probes for the maize microarrays were generated using the Fluorescent Linear Amplification Kit (cat. No. G2556A) from Agilent Technologies (Palo Alto, CA).

#### HYBRIDIZATION AND WASH CONDITIONS

The following Hybridization and Washing Condition were developed:

##### Hybridization Conditions:

Labeled probe was heated at 95°C for 3 min and chilled on ice. Then 25 µl of the hybridization buffer which was warmed at 42°C was added to the probe, mixing by pipetting, to give a final concentration of:

50% formamide

4x SSC

0.03% SDS

5x Denhardt's solution

0.1 µg/ml single-stranded salmon sperm DNA

5

The probe was kept at 42°C. Prior to the hybridization, the probe was heated for 1 more min., added to the array, and then covered with a glass cover slip. Slides were placed in hybridization chambers (Telechem, Sunnyvale, California) and incubated at 42°C overnight.

10

Washing Conditions:

- A. Slides were washed in 1x SSC + 0.03% SDS solution at room temperature for 5 minutes,
- B. Slides were washed in 0.2x SSC at room temperature for 5 minutes,
- C. Slides were washed in 0.05x SSC at room temperature for 5 minutes.

15

After A, B, and C, slides were spun at 800 x g for 2 min. to dry. They were then scanned.

20

Maize microarrays were hybridized according to the instructions included Fluorescent Linear Amplification Kit (cat. No. G2556A) from Agilent Technologies (Palo Alto, CA).

SCANNING OF SLIDES

The chips were scanned using a ScanArray 3000 or 5000 (General Scanning, Watertown, Massachusetts, USA). The chips were scanned at 543 and 633nm, at 10 µm resolution to measure the intensity of the two fluorescent dyes incorporated into the samples hybridized to the chips.

25

DATA EXTRACTION AND ANALYSIS

The images generated by scanning slides consisted of two 16-bit TIFF images representing the fluorescent emissions of the two samples at each arrayed spot. These images were then quantified and processed for expression analysis using the data extraction software Imagene™ (Biodiscovery, Los Angeles, California, USA). Imagene output was subsequently analyzed using the analysis program Genespring™ (Silicon Genetics, San Carlos, California,

30

USA). In Genespring, the data was imported using median pixel intensity measurements derived from Imagen output. Background subtraction, ratio calculation and normalization were all conducted in Genespring. Normalization was achieved by breaking the data in to 32 groups, each of which represented one of the 32 pin printing regions on the microarray. Groups consist of 360 to 550 spots. Each group was independently normalized by setting the median of ratios to one and multiplying ratios by the appropriate factor.

#### EXAMPLE 4: AFLP EXPERIMENTS AND RESULTS

##### Production of Samples

mRNA was prepared from 27 plant tissues. Based on preliminary cDNA-AFLP analysis with a few primer combinations, 11 plant tissues and/or pooled samples were selected. Samples were selected to give the greatest representation of unique band upon electrophoresis. The final 11 samples or pooled samples used in the cDNA-AFLP analysis were:

- |     |   |
|-----|---|
| S1  | Dark adapted seedlings                      |
| S2  | Roots/Etiolated Seedlings                   |
| S3  | Mature leaves, soil grown                   |
| S4  | Immature buds, inflorescence meristem       |
| S5  | Flowers opened                              |
| S6  | Siliques, all stages                        |
| S7  | Senescing leaves (just beginning to yellow) |
| S8  | Callus Inducing medium                      |
|     | Callus shoot induction                      |
|     | Callus root induction                       |
| S9  | Wounding                                    |
|     | Methyl-jasmonate-treated                    |
| S10 | Oxidative stress                            |
|     | Drought stress                              |
|     | Oxygen Stress-flooding                      |
| S11 | Heat treated light grown seedling           |

## Cold treated light grown seedlings

cDNA from each of the 11 samples was digested with two restriction endonucleases, namely *TaqI* and *MseI*. *TaqI* and *MseI* adapters were then ligated to the restriction enzyme  
5 fragments. Using primers to these adapters that were specific in sequence (i.e. without extensions), the restriction fragments were subjected to cycles of non-radioactive pre-amplification.

Selective PCR

10 In order to limit the number of fragments or bands on each lane of the AFLP gel, fragments were subjected to another round of selective radioactive polymerase chain amplification. The *TaqI* primers used in this amplification were 5'-labelled with P<sup>33</sup>. For these amplifications, the *TaqI* primers had two extra nucleotides at their 3' end and the *MseI* primers had three extra nucleotides at their 3' end. This resulted in 16 primer designs for the *TaqI* primer and 64 primer designs for the *MseI* primer. Altogether, this gave rise to a total of 1024 primer designs. Fragments generated in this selective amplification protocol were run with labeled  
15 molecular weight markers on polyacrylamide gels to separate fragments in the size range of 100 – 600 nucleotides.

Following gel electrophoresis, profiles were analyzed with a phosphoimager. From these  
20 images, electronic files, giving the mobilities of all bands on the gels and their intensities in each of the samples, were compiled.

All unique bands were cut out of the gels. The gel pieces were placed in 96 well plates for elution and their plate designation was linked to their electrophoretic mobilities recorded in the electronic files. The eluted fragments were then subjected to another round of amplification,  
25 this time using reamplification primers (see below). After amplification, DNA fragments were sequenced.

A computer database was established linking the mobilities of all the bands observed on the cDNA-AFLP gels with the sequence of the correspondingly isolated fragment. The sequence allowed for identification of the gene from which the cDNA-AFLP fragment was derived,  
30 allowing for a linkage of band mobility with the transcript of a specific gene. Also linked to the

This cDNA-AFLP analysis with *TaqI/MseI* and 1024 primer combinations was repeated using the enzymes *NlaIII* in place of *TaqI*, and *Csp6I* in place of *MseI*.

Experimental Samples were subjected to cDNA-AFLP as described above, resulting in electronic files recording band mobilities and intensities. Through use of the database established above, band mobilities could be linked to specific cDNAs, and therefore genes.

### Reamplification primers

CGCCAGGGTTTCCAGTCACGAC|ACGACTCACT|gatgagtcctgagtaa|

99G20

AGCGGATAACAATTCACACAGGA|CACACTGGTA|tagactgcgtaccga|

M13 reverse +10 TaqI +0

### Purification of the Reamplification reaction before sequencing

5  $\mu$ l reamplification reaction

0,25 µl 10xPCR buffer

0,33 µl Shrimp Alkaline Phosphatase (Amersham Life Science)

0,033 µl Exonuclease I (USB)

0,297 µl SAP dilution buffer

1,59  $\mu$ l MQ7.5  $\mu$ l total



30' 37°C

10' 80°C

4°C

#### Sample Preparation

S1: Dark adapted seedlings: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. After 8 days, the seedlings were foil-wrapped and harvested after two days.

S2: Roots/Etiolated seedlings: Seeds of *Arabidopsis thaliana* (wassilewskija) were germinated on solid germination media (1X MS salts, 1X MS vitamins, 20g/L sucrose, 50 mg/L MES pH 5.8) in the dark. Tissues were harvested 14 days later.

S3: Mature leaves, soil grown: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. Leaves were harvested 17 days later from plants that had not yet bolted.

S4: Immature buds, inflorescence meristem: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark.

S5: Flowers, opened: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark.

S6: Siliques, all stages: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark.

S7: Senescing leaves (just beginning to yellow): Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. When the plant had leaves that were less than 50% yellow, the leaves that were just beginning to yellow were harvested.

S8:

Callus Inducing Medium: Seeds of *Arabidopsis thaliana* (wassilewskija) were surface sterilized (1 min-75% Ethanol, 6 min-bleach 100% + Tween 20, rinse) and incubated on MS medium containing 2,4-Dichlorophenoxyacetic acid (2,4-D) 1 mg/l and Kinetin 1 mg/l in the dark for 3 weeks to generate primary callus.

Hypocotyls and roots of the seedling were swollen after a week after incubation in this callus induction medium and subsequently callus was initiated from these swollen areas.

Callus shoot induction: Primary calluses were transferred to the fresh callus induction medium for another 2 weeks growth to generate secondary callus. Secondary callus were transferred to shoot induction medium containing MS basal medium and Benzyladenine (BA) 2 mg/l and Naphthaleneacetic acid (NAA) .1 mg/l for 2 weeks growth in the light before it was harvested and frozen and sent to Keygene. Many shoot meristems were observed under the microscope.

Callus root induction: Secondary calluses were transferred to root induction medium containing MS basal medium, sucrose 1% and Indolebutyric acid (IBA) 0.05 mg/l in the dark. Many root primordia were observed under microscope after 10 days in the root induction medium. Those callus tissue were harvested and frozen and sent to Keygene.

S9:

Wounding: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. After 20 days, leaves of plants were wounded with pliers. Wounded leaves were harvested 1 hour and 4 hours after wounding.

Methyl jasmonate treatment: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. After 13 days, plants were sprayed with 0.001% methyl jasmonate. Leaves were harvested 1.5 hours and 6 hours after spraying

S10:

Oxidative stress: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. After 24 days, a few leaves were inoculated with a mixture of 2.5 mM D-glucose, 2.5 U/mL glucose oxidase in 20 mM sodium phosphate buffer pH 6.5. After an hour, 3 hours, or 5 hours after inoculation, whole plant, except for the inoculated leaves, was harvested. This sample was mixed with sample from plants that were sitting in full sun (152,000 LUX) for 2 hours or four hours.

Drought stress: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. After 20 days, aerial tissues were harvested and left to dry in 3MM Whatman paper for 1 hour or 4 hours.

Oxygen stress: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C,

with 16 h light and 8 h dark. After 21 days, the plant was flooded by immersing its pot in a beaker of tap water. After 6 days, the upper tissues were harvested.

S11: Heat-treated light grown seedlings: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. Over a 5 hour period, the temperature was raised to 42°C at the rate of approximately 4°C per hour. After 1 hour at 42°C, the aerial tissues were collected. This sample was mixed with an equal volume of sample that went through a heat-recovery treatment namely bringing down the temperature to 22°C from 42°C over a 5 hour period at the rate of 4°C per hour.

Cold-treated light grown seedlings: Seeds of *Arabidopsis thaliana* (wassilewskija) were sown in pots and left at 4°C for two to three days to vernalize. They were transferred to a growth chamber after three days. The intensity of light in the growth chamber was 7000-8000 LUX, temperature was 22°C, with 16 h light and 8 h dark. After 18 days, the plant was transferred to 4°C for an hour before the aerial tissues were harvested. This sample was mixed with aerial tissues from another plant that was transferred to 4°C for 27 hours before being harvested.

#### Analysis of Data:

Intensity: The intensity of the band corresponds to the value in each lane marked S1, S2 etc.

P-values: The data shows P- values of each of the samples 1-11. P-values are calculated using the following formula  $2*(1-NORMDIST(ABS(Sx-AVERAGE(of S1 to S11, not including Sx))/STDEV(of S1 to S11 not including Sx),0,1,TRUE))$  using Excel functions.

The equivalent mathematical formula of P-value is as follows:

$\int \phi(x) dx$ , integrated from  $a$  to  $\infty$ ,

where  $\phi(x)$  is a normal distribution:

where  $a = \frac{|Sx - \mu|}{\sigma}$

$\sigma(S1...S11, \text{ not including } Sx);$

where  $\mu$  = is the average of the intensities of all samples except  $Sx$ ,

$$= \frac{(\sum S1...Sn) - Sx}{n-1}$$

where  $\sigma(S1...S11, \text{ not including } Sx)$  = the standard deviation of all sample intensities except  $Sx$ .

#### Results:

The results are shown in the MA\_diff and/or AFLP\_diff AFLP\_int and AFLP\_diff tables.

#### EXAMPLE 5: TRANSFORMATION OF CARROT CELLS

Transformation of plant cells can be accomplished by a number of methods, as described above. Similarly, a number of plant genera can be regenerated from tissue culture following transformation. Transformation and regeneration of carrot cells as described herein is illustrative.

Single cell suspension cultures of carrot (*Daucus carota*) cells are established from hypocotyls of cultivar Early Nantes in B<sub>5</sub> growth medium (O.L. Gamborg et al., *Plant Physiol.* 45:372 (1970)) plus 2,4-D and 15 mM CaCl<sub>2</sub> (B<sub>5</sub> -44 medium) by methods known in the art. The suspension cultures are subcultured by adding 10 ml of the suspension culture to 40 ml of B<sub>5</sub>-44 medium in 250 ml flasks every 7 days and are maintained in a shaker at 150 rpm at 27 °C in the dark.

The suspension culture cells are transformed with exogenous DNA as described by Z. Chen et al. *Plant Mol. Bio.* 36:163 (1998). Briefly, 4-days post-subculture cells are incubated with cell wall digestion solution containing 0.4 M sorbitol, 2% driselase, 5mM MES (2-[N-Morpholino] ethanesulfonic acid) pH 5.0 for 5 hours. The digested cells are pelleted gently at 60 xg for 5 min. and washed twice in W5 solution containing 154 mM NaCl, 5 mM KCl, 125 mM CaCl<sub>2</sub> and 5mM glucose, pH 6.0. The protoplasts are suspended in MC solution containing 5 mM MES, 20 mM CaCl<sub>2</sub>, 0.5 M mannitol, pH 5.7 and the protoplast density is adjusted to about  $4 \times 10^6$  protoplasts per ml.

15-60  $\mu$ g of plasmid DNA is mixed with 0.9 ml of protoplasts. The resulting suspension is mixed with 40% polyethylene glycol (MW 8000, PEG 8000), by gentle inversion a few times

at room temperature for 5 to 25 min. Protoplast culture medium known in the art is added into the PEG-DNA-protoplast mixture. Protoplasts are incubated in the culture medium for 24 hour to 5 days and cell extracts can be used for assay of transient expression of the introduced gene. Alternatively, transformed cells can be used to produce transgenic callus, which in turn can be used to produce transgenic plants, by methods known in the art. See, for example, Nomura and Komamine, *Plt. Phys.* 79:988-991 (1985), *Identification and Isolation of Single Cells that Produce Somatic Embryos in Carrot Suspension Cultures*.

#### EXAMPLE 6: PHENOTYPE SCREENS AND RESULTS

##### A: Triparental Mating and Vacuum Infiltration Transformation of Plants

Standard laboratory techniques are as described in Sambrook et al. (1989) unless otherwise stated. Single colonies of *Agrobacterium* C58C1Rif, *E. coli* helper strain HB101 and the *E. coli* strain containing the transformation construct to be mobilized into *Agrobacterium* were separately inoculated into appropriate growth media and stationary cultures produced. 100 µl of each of the three cultures were mixed gently, plated on YEB (5g Gibco beef extract, 1g Bacto yeast extract, 1g Bacto peptone, 5g sucrose, pH 7.4) solid growth media and incubated overnight at 28°C. The bacteria from the triparental mating were collected in 2 ml of lambda buffer (20 mM Tris (pH 7.5), 100 mM NaCl, 10 mM MgCl<sub>2</sub>) and serial dilutions made. An aliquot of the each dilution was then plated and incubated for 2 days at 28°C on YEB plates supplemented with 100 µg/ml rifampicin and 100 µg/ml carbenicillin for calculation of the number of acceptor cells and on YEB plates supplemented with 100 µg/ml rifampicin, 100 µg/ml carbenicillin and 100 µg/ml spectinomycin for selection of transconjugant cells. The cointegrate structure of purified transconjugants was verified via Southern blot hybridization.

A transconjugant culture was prepared for vacuum infiltration by inoculating 1 ml of a stationary culture arising from a single colony into liquid YEB media and incubating at 28°C for approximately 20 hours with shaking (220 rpm) until the OD taken at 600 nm was 0.8-1.0. The culture was then pelleted (8000 rpm, 10 min, 4°C in a Sorvall SLA 3000 rotor) and the bacteria resuspended in infiltration medium (0.5X MS salts, 5% w/v sucrose, 10 µg/l BAP, 200 µl/l Silwet L-77, pH 5.8) to a final OD<sub>600</sub> of 1.0. This prepared transconjugant culture was used within 20 minutes of preparation.

Wild-type plants for vacuum infiltration were grown in 4-inch pots containing Metromix 200 and Osmocote. Briefly, seeds of *Arabidopsis thaliana* (ecotype Wassilewskija) were sown in pots and left at 4°C for two to four days to vernalize. They were then transferred to 22-25°C and grown under long-day (16 hr light: 8 hr dark) conditions, sub-irrigated with water. After bolting, the primary inflorescence was removed and, after four to eight days, the pots containing the plants were inverted in the vacuum chamber to submerge all of the plants in the prepared transconjugant culture. Vacuum was drawn for two minutes before pots were removed, covered with plastic wrap and incubated in a cool room under darkness or very low light for one to two days. The plastic wrap was then removed, the plants returned to their previous growing conditions and subsequently produced (T1) seed collected.

#### B: Selection of T-DNA Insertion Lines

Approximately 10,750 seeds from the initial vacuum infiltrated plants were sown per flat of Metromix 350 soil. Flats were vernalized for four to five days at 4°C before being transferred to

22-25°C and grown under long-day (16 hr light: 8 hr dark) conditions, sub-irrigated with water. Approximately seven to ten days after germination, the (T1) seedlings were sprayed with 0.02% Finale herbicide (AgrEvo). After another five to seven days, herbicide treatment was repeated. Herbicide resistant T1 plants were allowed to self-pollinate and T2 seed were collected from each individual. In the few cases where the T1 plant produced few seed, the T2 seed was planted in bulk, the T2 plants allowed to self-pollinate and T3 seed collected.

#### C: Phenotype Screening

Approximately 40 seed from each T2 (or T3) line were planted in a 4-inch pot containing either Sunshine mix or Metromix 350 soil. Pots were vernalized for four to five days at 4°C before being transferred to 22-25°C and grown under long-day (16 hr light: 8 hr dark) conditions, sub-irrigated with water. A first phenotype screen was conducted by visually inspecting the seedlings five to seven days after germination and aberrant phenotypes noted. Plants were then sprayed with Finale herbicide within four days (i.e. about seven to nine days after germination).

The second visual screen was conducted on surviving T2 (or T3) plants about sixteen to seventeen days after germination and the final screen was conducted after the plants had bolted and formed siliques. Here, the third and fourth green siliques were collected and aberrant phenotypes noted. The Knock-in and Knock-out Tables contain descriptions of identified phenotypes.

Alternative, seed were surface sterilized and transferred to agar solidified medium containing Murashige and Skoog salts (1X), 1% sucrose (wt/v) pH 5.7 before autoclaving. Seed were cold treated for 48 hours and transferred to long days [16 hours light and 8 hours dark], 25°C. Plants were screened at 5 and 10 days.

In another screen, seed were surface sterilized and transferred to agar solidified medium containing Murashige and Skoog salts (1X), and combinations of various nitrogen and sucrose amounts as specified below::

Medium 1: no sucrose, 20.6 mM  $\text{NH}_4\text{NO}_3$ , 18.8 mM  $\text{KNO}_3$ ;

Medium 2: 0.5% sucrose, 20.6 mM  $\text{NH}_4\text{NO}_3$ , 18.8 mM  $\text{KNO}_3$ ;

Medium 3: 3% sucrose, 20.6 mM  $\text{NH}_4\text{NO}_3$ , 18.8 mM  $\text{KNO}_3$ ;

Medium 4: no sucrose, 20.6  $\mu\text{M}$   $\text{NH}_4\text{NO}_3$ , 18.8  $\mu\text{M}$   $\text{KNO}_3$ ;

Medium 5: 0.5% sucrose, 20.6  $\mu\text{M}$   $\text{NH}_4\text{NO}_3$ , 18.8  $\mu\text{M}$   $\text{KNO}_3$ ; and

Medium 6: 3% sucrose, 20.6  $\mu\text{M}$   $\text{NH}_4\text{NO}_3$ , 18.8  $\mu\text{M}$   $\text{KNO}_3$ .

The 0.5% sucrose was the control concentration for the sucrose. The low nitrogen, 20.6  $\mu\text{M}$   $\text{NH}_4\text{NO}_3$ , 18.8  $\mu\text{M}$   $\text{KNO}_3$ , is the control for the nitrogen. Seed were cold treated for 48 hours and transferred to long days [16 hours light and 8 hours dark], 25°C. Plants were screened at 2, 5, and 10 days.

#### D: TAIL-PCR and Fragment Sequencing

Rosette leaves were collected from each putative mutant and crushed between parafilm and FTA paper (Life Technologies). Two  $2\text{mm}^2$  hole punches were isolated from each FTA sample and washed according to the manufacturer's instructions by vortexing with 200  $\mu\text{l}$  of the provided FTA purification reagent. The FTA reagent was removed and the washing procedure repeated two more times. The sample was then washed twice with 200  $\mu\text{l}$  of FTA TE (10 mM Tris, 0.1 mM EDTA, pH 8.0) and vortexing prior to PCR.



Primers used for TAIL-PCR are as follows:

AD2: 5' NGTCGASWGANAWGAA 3' (128-fold degeneracy)

S = G or C, W = A or T, and N = A, G, C, or T

LB1: 5' GTTTAACTGCGGCTCAACTGTCT 3'

LB2: 5' CCCATAGACCCCTTACCGCTTTAGTT 3'

LB3: 5' GAAAGAAAAAGAGGTATAACTGGTA 3'

The extent to which the left and right borders of the T-DNA insert were intact was measured for each line by PCR. The following components were mixed for PCR: 1 2mm<sup>2</sup> FTA sample, 38.75 µl distilled water, 5 µl 10X Platinum PCR buffer (Life Technologies), 2 µl 50 mM MgCl<sub>2</sub>, 1 µl 10 mM dNTPs, 1 µl 10 µM primer LB1 (or RB1 for analysis of the right border), 1 µl 10 µM primer LB3R (or RB3R for analysis of the right border) and 1.25 U Platinum Taq (Life Technologies). Cycling conditions were: 94°C, 10 sec.; thirty cycles of 94°C, 1 sec. - 54°C, 1 sec. - 72°C, 1 sec.; 72°C, 4 sec. The expected band size for an intact left border is bp, while an intact right border generates a bp band.

Fragments containing left or right border T-DNA sequence and adjacent genomic DNA sequence were obtained via PCR. First product PCR reactions use the following reaction mixture: 1 2mm<sup>2</sup> FTA sample, 12.44 µl distilled water, 2 µl 10X Platinum PCR buffer (Life Technologies), 0.6 µl 50 mM MgCl<sub>2</sub>, 0.4 µl 10 mM dNTPs, 0.4 µl 10 µM primer LB1 (or RB1 for analysis of the right border), 3 µl 20 µM primer AD2 and 0.8 U Platinum Taq (Life Technologies). Cycling conditions for these reactions were: 93°C, 1 min.; 95°C, 1 min.; three cycles of 94°C, 45 sec. - 62°C, 1 min. - 72°C, 2.5 min.; 94°C, 45 sec.; 25°C, 3 min.; ramp to 72°C in 3 min.; 72°C, 2.5 min.; fourteen cycles of 94°C, 20 sec. - 68°C, 1 min. - 72°C, 2.5 min. - 94°C, 20 sec.; - 68°C, 1 min. - 72°C, 2.5 min. - 94°C, 20 sec. - 44°C, 1 min. - 72°C, 2.5 min.; 72°C, 5 min.; end; ~4.5 hrs. For second product PCR reactions 1 µl of a 1:50 dilution of the first PCR product reaction was mixed with 13.44 µl distilled water, 2 µl 10X Platinum PCR buffer (Life

Technologies), 0.6  $\mu$ l 50 mM  $MgCl_2$ , 0.4  $\mu$ l 10 mM dNTPs, 0.4  $\mu$ l 10  $\mu$ M primer LB2 (or RB2 for analysis of the right border), 2  $\mu$ l 20  $\mu$ M primer AD2 and 0.8 U Platinum Taq (Life Technologies). Second product cycling conditions were: eleven cycles of 94°C, 20 sec. - 64°C, 1 min. - 72°C, 2.5 min. - 94°C, 20 sec. - 64°C, 1 min. - 72°C, 2.5 min. - 94°C, 20 sec. - 44°C, 1 min.; 72°C, 5 min.; end; ~3 hrs. Third product PCR reactions were prepared by first diluting 2  $\mu$ l of the second PCR product with 98  $\mu$ l of distilled water and then adding 1  $\mu$ l of the dilution to 13.44  $\mu$ l distilled water, 2  $\mu$ l 10X Platinum PCR buffer (Life Technologies), 0.6  $\mu$ l 50 mM  $MgCl_2$ , 0.4  $\mu$ l 10 mM dNTPs, 0.4  $\mu$ l 10  $\mu$ M primer LB3 (or RB3 for analysis of the right border), 2  $\mu$ l 20  $\mu$ M primer AD2 and 0.8 U Platinum Taq (Life Technologies). Third product cycling conditions were: twenty cycles of 94°C, 38 sec. - 44°C, 1 min. - 72°C, 2.5 min.; 72°C, 5 min.; end; ~2 hrs. Aliquots of the first, second and third PCR products were electrophoresed on 1% TAE (40 mM Tris-acetate, 1 mM EDTA) to determine their size.

Reactions were purified prior to sequencing by conducting a final PCR reaction. Here, 0.25  $\mu$ l Platinum PCR Buffer (Life Technologies), 0.1  $\mu$ l 50 mM  $MgCl_2$ , 3.3 U SAP shrimp alkaline phosphatase, 0.33 U Exonuclease and 1.781  $\mu$ l distilled water were added to a 5  $\mu$ l third product and the reaction cycled at 37°C, 30 min.; 80°C, 10 min.; 4°C indefinitely.

Di-deoxy "Big Dye" sequencing was conducted on Perkin-Elmer 3700 or 377 machines.

### KNOCK-IN EXPERIMENTS

For the following examples, a two-component system was constructed in a plant to ectopically express the desired cDNA.

First, a plant was generated by inserting a sequence encoding a transcriptional activator downstream of a desired promoter, thereby creating a first component where the desired promoter facilitates expression of the activator generated a plant. The first component also is referred to as the activator line.

Next, the second component is constructed by linking a desired cDNA to a sequence that the transcriptional activator can bind to and facilitate expression of the desired cDNA. The second component can be inserted into the activator line by transformation. Alternatively, the

second component can be inserted into a separate plant, also referred to as the target line. Then, the target and activator lines can be crossed to generate progeny that have both components.

Two component lines were generated by both means.

#### Part I - From crosses

Target lines containing cDNA constructs are generated using the Agrobacterium-mediated transformation. Selected target lines are genetically crossed to activation lines (or promoter lines). Generally, the promoter lines used are as described above. Evaluation of phenotypes is done on the resulting F1 progenies.

#### Part II - From Type I Supertransformation

Promoter activation lines (generally Vascular/Ovule/Young Seed/Embryo line, Seed/Epidermis/Ovary/Fruit line, Roots/Shoots/Ovule line, and Vasculature/Meristem are transformed with cDNA constructs using the Agrobacterium mediated transformation. Selected transformants (and their progenies) are evaluated for changes in phenotypes. The table for the knock-in of the Type I supertransformation comprises the following information

- Clone ID,
- Pfam,
- Gemini ID
- Trans. Unique ID (which indicates what promoter activation line was transformed
- S Ratio: segregation ratio after the transformed plants are selected for the marker.
- Assay
- Stage: phenotype was observed
- Feature: Where the phenotype was observed
- Phenotype

- P Ratio: phenotype ratio
- Comments

5 Part III - From Type II Supertransformation

Target lines generated using the procedure mentioned in Part I are transformed with T-DNA construct containing constitutive promoter. Selected transformants (and their progenies) are evaluated for changes in phenotypes.

10

An additional deposit of an *E. coli* Library, *E. coli*LibA021800, was made at the American Type Culture Collection in Manassas, Virginia, USA on February 22, 2000 to meet the requirements of Budapest Treaty for the international recognition of the deposit of microorganisms. This deposit was assigned ATCC accession no. PTA-1411.

Additionally, ATCC Library deposits; PTA-1161, PTA-1411 and PTA-2007 were made at the American Type Culture Collection in Manassas, Virginia, USA on; January 7, 2000, February 23, 2000 and June 8, 2000 respectively, to meet the requirements of Budapest Treaty for the international recognition of the deposit of microorganisms.

The invention being thus described, it will be apparent to one of ordinary skill in the art that various modifications of the materials and methods for practicing the invention can be made. Such modifications are to be considered within the scope of the invention as defined by the following claims.

Each of the references from the patent and periodical literature cited herein is hereby expressly incorporated in its entirety by such citation.

25